Journal of Experimental Psychology

THE JOHN CREHAR LIBRARY
RECEIVED 1952

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

The Role of Physiological Nystagmus in Monocular Acuity: F. RATLIFF 163

Experiential Factors in Visual Form Perception: II. Latency as a Function of Repetition: R. C. ATKINSON AND R. B. AMMONS 173

The Effect of Simultaneous Visual Stimulation of Absolute Auditory Sensitivity: L. W. GREGG AND W. J. BROGDEN 179

The Perception of Obstacles by the Deaf: P. WORCHEL AND J. H. BERRY 187

Frequency of Usage as a Determinant of Recognition Thresholds for Words: R. L. SOLOMON AND L. POSTMAN 195

Errors of Visual Recognition: F. H. GEORGE 202

Retention of Verbal Material as a Function of Motivating Instructions and Experimentally-Induced Failure: W. A. RUSSELL 207

Displacement: Greater Generalization of Approach than Avoidance in a Generalized Approach-Avoidance Conflict: N. E. MILLER AND D. KRAELING 217

Displacement: Steeper Gradient of Generalization of Avoidance than of Approach with Age of Habit Controlled: E. J. MURRAY AND N. E. MILLER 222

Displacement and Conflict: Learnable Drive as a Basis for the Steeper Gradient of Avoidance than of Approach: N. E. MILLER AND E. J. MURRAY 227

An Equal Discriminability Scale for Loudness Judgments: W. R. GARNER 232

Accuracy of Tactual Discrimination of Letters, Numerals, and Geometric Forms: T. R. AUSTIN AND R. B. SLEIGHT 239

American Psychological Association

Vol. 43 No. 3 March 1952

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.00, or $14.00 annually. Single copies are $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., Prince and Lemon Sts., Lancaster, Pa., or 1515 Massachusetts Avenue, N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental psychology or for behavior theory. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, a single experimental study or an integrated series of studies accomplished simultaneously must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free reprints without covers. However, under the new APA reprint policy no free reprints will be supplied to authors submitting manuscripts after January 1, 1952. All of the cost of the author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas.

Immediate publication can be given to articles whose authors assume the full cost of publication. Such articles must undergo the usual review by the Consulting Editors of the JOURNAL. No free reprints are given in the case of these articles.

Address all articles submitted for regular publication or for immediate publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

The JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. Anderson and W. L. Valentine, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *E*s, *E*'s, and *E*s'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts must be typed double-spaced throughout, including tables, footnotes, references, and titles for figures. Tables, figures, footnotes, references, and titles for figures must be placed on separate pages at the end of the manuscript.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 43, NO. 3 MARCH, 1952

THE ROLE OF PHYSIOLOGICAL NYSTAGMUS IN MONOCULAR ACUITY [1]

FLOYD RATLIFF [2]
The Johns Hopkins University

The extremely delicate acuity of the human visual system and its ability to perceive sharp and clearly defined contours are two of the most perplexing phenomena observed in the study of visual processes. In the first place, the limits of visual acuity appear to exceed physiological limits in that certain minimum perceptible visual angles of test objects are smaller than the visual angles subtended by single retinal receptors. Secondly, the perception of contour is clear and sharp even though the retinal image is blurred to a certain degree (because of diffraction and aberrations) and the pattern of the retinal mosaic is somewhat irregular. These facts are not adequately explained by the widely accepted classical theories of visual function. There are several modern theories of visual acuity, however, which postulate the existence of various mechanisms presumed to be capable of overcoming these limiting factors.

One common element in many of these more recent theories (**4, 6, 12**) is the notion that the minute involuntary tremor of the eye (physiological nystagmus) plays a part in the detection of extremely small separations and vernier offsets in test objects and in the sharpening of the blurred contours of the retinal images. In general, the tremor is thought of as (a) the basis for an averaging mechanism which somehow "rectifies" the blurred image on the retina, or (b) the means by which the illumination on a given receptor is rapidly changed from one level to another, presumably resulting in more effective stimulation of certain receptor units. There is very little direct experimental evidence, however, to support these concepts. The primary purpose of the present experiments is to determine the role of physiological nystagmus in monocular acuity.

METHOD

The experiments consist, essentially, of obtaining records of the involuntary eye tremor during the presentation of an acuity test object,

[1] This paper is based on the experimental portion of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Brown University. The research was carried out under the direction of Dr. L. A. Riggs, and is part of Project NR-140-359, Contract N7 onr-358, Task Order II, between the Office of Naval Research and Brown University.

[2] Now at Harvard University.

and of measuring certain aspects of the records to determine whether or not they vary in any systematic way with the changes in the acuity judgments made by *S*.

A photographic method is used to record the physiological nystagmus. The method is essentially the same as the one described in an earlier paper (**8**). The basic element of the system is an optical lever provided by light reflected from a plane first-surface mirror mounted in a contact lens worn by *S*.

The presentation of the acuity test objects is accomplished by means of a tachistoscopic device in which a half-silvered mirror is utilized to obtain an almost instantaneous shift from the fixation field to the test-object field.

The *S* fixates monocularly upon crosshairs in the fixation field. A warning signal is given and about 1 sec. later the crosshairs and fixation field are replaced for a period of 75 msec. by a parallel bar acuity test object, after which the fixation field is restored. The test object is presented randomly in horizontal and vertical positions and *S*'s task is to report on the orientation of the parallel bars. In Fig. 1 is shown a diagram representing the appearance of the various fields from *S*'s point of view. During this time the eye movements are recorded by means of the mirror in the contact lens worn by *S*.

Apparatus for Presentation of Test Objects

The presentation of test objects is accomplished by means of the apparatus shown in Fig. 2. A rotary shutter is held in a stationary position by a detent which engages a notch in a cam attached to the shutter. When this detent is disengaged by tripping the clutch lever, a friction clutch is free to move the shutter. The rotary shutter almost immediately attains the speed of the rotating shaft, that is, one revolution per second. When the shutter completes one revolution, the detent again engages the notch in the cam, thus stopping the rotation of the shutter until the clutch lever is tripped again.

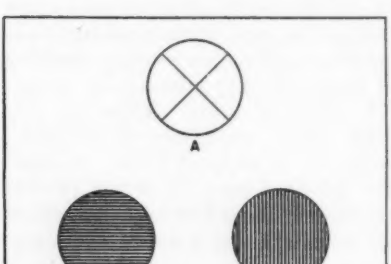

FIG. 1. The appearance of the crosshairs and fixation field (*A*), a horizontal test object (*B*), and a vertical test object (*C*)

The rotary shutter is constructed with an open sector, opposite which is a projecting vane. The vane and the open sector subtend angles (with vertices at the center of the shutter) which are exactly equal. Two projectors are mounted behind the shutter so that the plane of the shutter coincides with the focal plane of the projectors. When the shutter is in the stationary position, light from $Projector_1$ is allowed to pass by the shutter; and by means of a lens (L_1) and a mirror (M_1) it is made to fill a second lens (L_2). In front of this lens is placed a set of crosshairs (*CH*) made of fine human hair. These hairs, at 90° to one another, are placed at angles of 45° from the horizontal, with the crossed portion over the center of the lens. The crosshairs are in the focal plane of a third lens (L_3) so that rays from any point on the crosshairs are rendered parallel upon passing through lens L_3. These parallel rays pass through a half-silvered mirror (M_2) and are brought to a focus by another lens (L_5). Finally, another lens (L_6) which serves as the eyepiece for *S* renders the rays parallel again. These rays are then brought to a focus on the retina by the optical system of the eye (including the contact lens). *As a result of this arrangement, when the shutter is stationary, the* S *perceives the crosshairs surrounded by a circular homogeneously illuminated field* (as shown in Fig. 1A). The diameter of this field subtends a visual angle of approximately 9°.

When the clutch lever is tripped, the shutter is free to rotate through one complete revolution. During the time that the shutter is open for $Projector_2$ (and, consequently, closed for $Projector_1$) light from $Projector_2$ falls on a diffusing glass (*DG*). The parallel bar acuity test object (*TO*) abuts against this diffusing glass. The test object is in the focal plane of a lens (L_4) so that rays from any point on the test object are rendered parallel upon passing through this lens. Next, the parallel rays strike the half-silvered mirror (M_2) and are reflected, with some light loss, to lens L_5. This lens brings the rays to a focus and they are again rendered parallel by lens L_6. Again, these rays continue parallel until they are focused on the retina by the contact lens and refracting media of the eye. As a result of this arrangement, *during the period of 75 msec. when the crosshairs are not illuminated and the test object is illuminated,* S *perceives the parallel bar acuity test object in place of the crosshairs* (as in Fig. 1B or 1C). The crosshairs and homogeneous fixation field are then restored and remain visible until the clutch lever is tripped

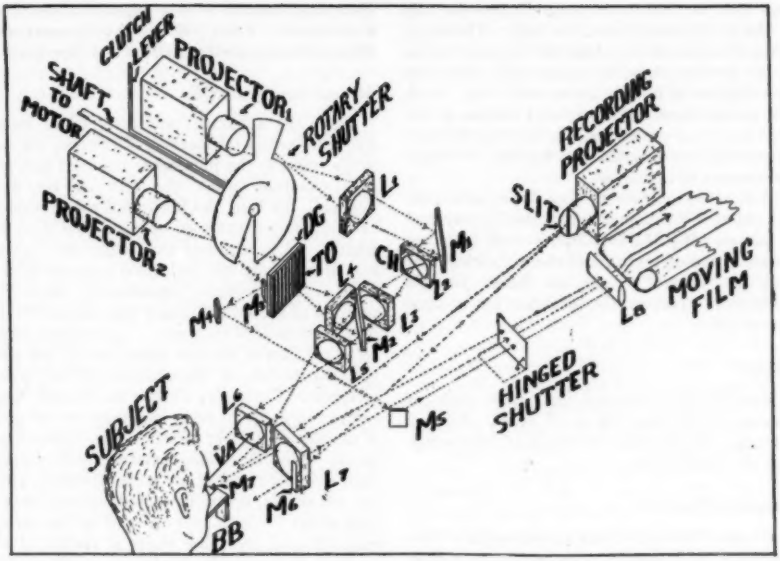

FIG. 2. The apparatus for presentation of the acuity test object and recording of the horizontal component of the physiological nystagmus

and the cycle is repeated. Since the rays striking the eye from any point on the crosshairs or test objects are parallel, the crosshairs and test objects are optically at an infinite distance from *S*. The brightness of the fixation field and that of the test-object field were equated so that there would be no flicker when the test objects were presented. The brightness of these fields, as measured by the Macbeth illuminometer, was approximately 325 ml.

Several test objects serving as measures of various degrees of visual acuity were available for use. These test objects consisted of alternate light and dark parallel lines of equal width. They were on 2 × 2-in. plates processed photographically and covered by ordinary lantern slide cover glasses to protect the emulsion. The visual angles subtended by the bars of the various test objects ranged from .98 min. to 1.64 min.

Apparatus for Photographic Recording

Each *S* was fitted with a contact lens for the right eye.[3] The lens was ground to correct for any correctable defects in *S*'s eye. By means of an end milling tool a hole approximately 4 mm. in diameter was drilled partially through each lens on the temporal side of the corneal bulge. A high quality first-surface plane mirror was turned down to a diameter slightly smaller than that of the hole in the contact lens. The mirror was then fitted tightly into the hole and cemented in place by means of "airplane glue." Soft wax was placed around the edge of the mirror to prevent irritation of *S*'s lids. A plastic contact lens prepared in this manner may be worn without discomfort for a period of an hour or more.

The source of light for the recording system (see Fig. 2) consists of a narrow vertical slit illuminated by a projector operated on direct current. The main lens (L_7) of the recording system has a focal length of one meter. This lens is in a position such that the illuminated slit is in its focal plane. This lens and the projector and slit are all fastened securely to a very heavy steel I-beam. A biting board (*BB*) for *S* is also attached to this beam. The arrangement is such that the rays from any point at the illuminated slit are rendered parallel upon passing through the main lens. The parallel rays strike the mirror (M_7) in the contact lens worn by *S*, and are reflected back through the main lens. (A straight-edged stop prevents any of these rays from reaching the pupil of *S*'s eye.) After being reflected back through the main lens, the

[3] I am indebted to Dr. E. V. Troendle for his skill in fitting these lenses.

rays come to a focus on a vertically moving strip of film in the focal plane of the lens. The image of the illuminated slit is brought to a point focus, in the vertical plane, by means of a cylindrical lens (L_6) just in front of the moving film. With this arrangement any horizontal motion of the eye results in a proportional horizontal displacement of the trace on the moving film. Vertical motions are not recorded.

A signal to indicate the time during which the test object is illuminated is provided by a system of mirrors (M_3, M_4, and M_5), a weak lens (not shown in the diagram), and the cylindrical lens (L_6). A "control" trace, to detect possible mechanical artifacts, is provided by a fixed mirror (M_6).

Subjects

Records were obtained from three male *S*s ranging in age from 24 to 37 years. All had "normal" (20/20 or better) Snellen Chart acuity with the unaided eye.

Recording Procedure

A pair of the test objects was mounted so that they could be presented either in a horizontal or vertical position. A prearranged randomized order (with respect to the orientation of the test object) of from 50 to 100 presentations was followed. The *S*, his head positioned by the biting board, fixated upon the crosshairs with his right eye. The left eye was occluded. After a few minutes of adaptation to the brightness level of the field, the recording was begun. Previously, *S* had been adapted to the level of the room illumination.

The *E*, by observing the position of the reflected beams on the face of the hinged shutter, could determine when the beam reflected from the mirror in the contact lens was in the proper position for recording. The *S* was given a warning signal and the clutch lever was tripped about .5 sec. later, resulting in exposure of the test object for a period of 75 msec. The *S* then judged the orientation of the parallel bars of the test object. Five categories of judgment were allowed: "vertical"; "horizontal"; "vertical-doubtful"; "horizontal-doubtful"; and "don't know." The first and second of these indicated that *S* was virtually certain that his judgment was correct. The third and fourth indicated some degree of uncertainty. The fifth indicated that he was absolutely uncertain as to the orientation of the test object. The *E* recorded the judgment made by *S* along with the true orientation of the test object.

Test objects were presented as rapidly as convenient, usually about every 30 sec. The *S* was allowed to rest at any time he wished to do so. Each experimental session required about 1 hr. to complete. From 200 to 500 judgments were obtained for each of the pairs of test objects used.

Measurement of Records

Since the apparatus was designed to record only the horizontal motions of the eye, use was made only of the judgments reported and records obtained when the test objects were vertical. In fact, the horizontal orientation of the test object was used only to keep *S* on his mettle. It is assumed that with the vertical test objects only the horizontal component of a given eye movement significantly alters the position of the retinal image with respect to the receptor mosaic of the fovea. Movements parallel to the bars of the test object would not alter the distribution of illumination on the foveal receptors. Certainly, no eye movements were sufficiently large to bring the image of the ends of the bars of the test objects to the region of the fovea. (No movements having extents greater than a few minutes of arc were recorded, while the test object, centrally fixated, extends over a field about 9° in diameter.) *All of the critical measurements, then, are those of motion of the receptors perpendicular to the bars of the test object, and not of motion parallel to these bars.*

Earlier investigations (8) have revealed that four "types" of involuntary motion of the eye occur during attempted fixation: (a) small rapid motions with a median extent (peak to trough) of about 17.5-sec. angle of rotation, and frequencies ranging from 30 to 70 cycles per second; (b) relatively slow motions of irregular

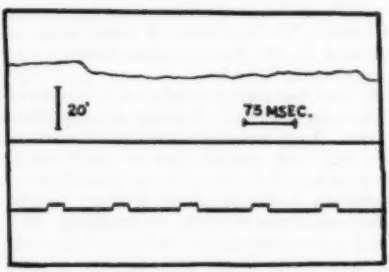

FIG. 3. A sample record of the horizontal component of the involuntary eye movements during steady fixation. Upper trace is the eye movement record. The trace in the center is a control trace provided by a fixed mirror. The time line (lower trace) is displaced for a short period every .1 sec. The vertical line drawn on the record subtends approximately 20 min. of arc. The horizontal line drawn on the record represents approximately 75 msec. of time.

frequency and amplitude; (c) slow drifts in one direction or another; and (d) rapid jerks with an average extent of about 5 or 6 min. of arc, occurring at irregular intervals, and at times apparently compensating for the drifts.

However, during a brief period such as the 75-msec. exposure time used in the present experiments it may be assumed for practical purposes that there are only two "types" of motions. The rapid jerky motions occur so rarely that they may be disregarded in these experiments. The slow waves and drifts may be lumped together since, during a period of 75 msec., the end result of either type of motion is a gradual change in the position of the eye in one direction or another. Therefore, the measurements to be discussed will be only the minute rapid tremor and the gradual drifts of the eye, the latter term including both the drifts and the relatively slow waves.

In the case of the recording apparatus used in these experiments it may be shown that the eye may be considered as rotating about a point at the center of the main lens of the recording system. The tangent of the angle of rotation may be computed by dividing the amount of displacement of the trace on the record by two times the distance from the main lens to the recording paper. (It is necessary to divide the displacement by two because the light beam reflected from a mirror turns through twice the angle of rotation of the mirror.) It has been shown earlier (8) that, if the fixation and test fields are optically at an infinite distance (as they are in this situation), the angle of rotation of the eye computed in this manner is equal to the visual angle subtended by the displacement of the retinal image. Therefore, all of the measurements to be discussed may be regarded as measurements of the motions of the retinal image.

Measurement of gradual drifts.—Since the extents of the drifts on the records are fairly small, it was necessary to use a microscope and suitable reticle to measure them.

The actual measurements of the drift consist of measuring the amount of displacement of the trace on the record from the beginning to the end of the stimulus signal. The measurement was arbitrarily made between corresponding points such as the peaks of the tremors nearest the ends of the signal. The control trace provided a base line from which the measurements could be made.

Measurement of minute tremor.—In order to avoid possible confusion of the effects of variations in the extents of the gradual drifts and the effects of variations in the minute tremor, only the tremor in records having approximately equal amount of drift was measured. It was arbitrarily decided to use only the records having relatively small amounts of drift (20 sec. of arc or less). Since the rapid tremor is extremely small, all of the records to be measured were enlarged about ten times to facilitate their measurement.

It was decided that a measure of the *average change in position* of the trace per unit time would best describe the variations in the minute tremor. The measurements were made by means of the following instrument. A needle-like pointer was attached to a movable shaft which was moved along the vertical axis of the record by means of a micrometer. A strong spring kept the shaft in firm contact with the micrometer at all times to prevent "lost motion." A handwheel attached to the micrometer made it possible to make very fine adjustments of the pointer. A mechanical counter attached to the handwheel recorded tenths of revolutions of the micrometer knob. One revolution moved the pointer a distance on the record equal to approximately 5 sec. of visual angle. A plate mounted on the cross slide of a lathe moved the record along the time axis with respect to the pointer.

The actual measurements of the tremor were made as follows. The record was mounted on the plate attached to the cross slide in a position such that the time axis of the record was parallel with the line of motion of the cross slide. The cross slide and pointer were then moved into positions such that the pointer was aligned with one edge of the trace at the beginning of the stimulus signal, and a reading of the counter was taken. The cross slide was then advanced by a given amount. (Successive advances of the cross slide have the effect of dividing the eye movement record into about 19 segments, each covering about 3.9 msec. of time.) The pointer was aligned each time with the edge of the trace by turning the micrometer handwheel, and a reading of the counter was taken at each of these successive positions. The resulting difference between successive readings represents the change in the position of the trace during the given interval of 3.9 msec. The successive differences were summed and divided by the number of time intervals to yield the *average change in position* of the eye per unit time during the exposure of the acuity test object.

In all of the measurements of drift and tremor described above, the records were identifiable only by means of a code number. The *E* did not know, until after the records were measured, which records corresponded to the various categories of judgment made by *S*.

Results

Once the measurements as described above were completed, they were related to the acuity judgments made at the time the records were obtained. Since only the records obtained when the acuity test objects were vertical were measured, the judgments made when the test objects were horizontal were discarded. It was found that few "doubtful" judgments were made, and that these judgments were always correct, as were the "vertical" judgments. Therefore, these two categories were lumped together and called "correct" responses. The remaining category, "don't know," was called "incorrect." It is important to note that horizontal test objects were presented randomly along with the vertical test objects and this served to eliminate guessing on the part of *S*. In fact, there was no evidence of guessing, such as judging a horizontal test object as vertical or vice versa.

Gradual Drift and Visual Acuity

The distribution of correct and incorrect responses with respect to the amount of drift suggested that when the drift is relatively small the proportion of correct responses is greater than when the drift is relatively large. In order to make a more careful analysis of this relationship, the drifts were divided into "small" drifts, which included those of 20 sec. of angle of rotation or less, and "large" drifts, which included those greater than 20 sec. of angle of rotation of the eye. This resulted in roughly equal frequencies in the two classes. It should be remembered, however, that even the extents of the very largest drifts (140 to 150 sec.) are only about twice the width of the bars of the acuity test objects near the threshold.

In Fig. 4 are shown graphs of some of the data obtained from each of the three *S*s, using test objects near their acuity thresholds. In each case a higher percentage of correct responses results when the drift is relatively small than when it is relatively large.

A statistical test was made of the hypothesis that the number of correct and incorrect responses is independent of the extent of the drift. Using a 2×2 χ^2 table (correct and incorrect responses, large and small drift) it was found that for subject LAR, p is less than .001, for LH, p is less than .001, and for FR, p is less than .01. Therefore, it may be concluded that relatively large extents of drift are a hindrance to monocular acuity as measured in these experiments.

Minute Tremor and Visual Acuity

Examination of the records and of the data obtained by measuring the records of the minute tremor of the eye suggested that the incorrect responses were associated with the larger amplitudes of the minute tre-

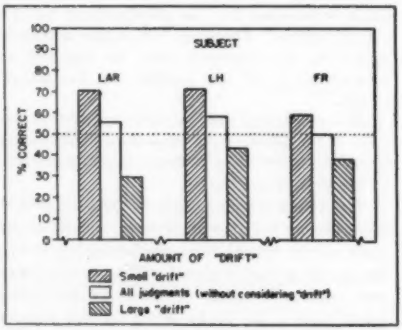

Fig. 4. Per cent of correct acuity judgments when the drift of the visual axis was "small," "large," and when drift was not considered. In each case the test object was near the subject's acuity threshold. The bars of the test objects subtended the following visual angles: for subject LAR, 1.077 min.; for subject LH, 1.025 min.; for subject FR, 1.64 min.

TABLE 1

MEAN CHANGE IN POSITION OF THE EYE (SEC.) PER UNIT TIME (3.9 MSEC.) FOR CORRECT AND INCORRECT JUDGMENTS

Subject	Judgment	*N*	Mean	*p*
LAR	Correct	35	2.81	<.02
	Incorrect	15	3.53	
LH	Correct	50	3.24	<.05
	Incorrect	20	3.83	
FR	Correct	50	2.92	<.40
	Incorrect	25	3.08	

mor, while the correct responses were associated with the smaller amplitudes. However, since the tremor is only one of a great number of factors determining the limits of acuity, some statistical treatment of the data was necessary to reveal the relationship more clearly. The "average change in position" of the eye per unit time was determined for each of the records as outlined above. Next, for each *S* the means of the average changes in the records pertaining to correct judgments and of those pertaining to incorrect judgments were determined. For each *S* it was found that the mean amount of tremor, as measured in this manner, was greater in the case of incorrect responses than in the case of correct responses.

Using a *t*-test for small independent samples the *p*-values shown in Table 1 were obtained.

Although the difference for subject FR is not significant, it should be noted that the trend is in the same direction as in the other two cases. Therefore, it may be concluded that relatively large changes in the position of the eye (as a result of the rapid tremor) are a hindrance to monocular acuity as measured in these experiments.

DISCUSSION

It has been pointed out in various reviews of theories of visual acuity (**10, 11**) that there are two general types of acuity theory. First, there are the so-called "static" theories based largely on peripheral processes. In these theories the resolving power of the eye is accounted for on the basis of the ability of the visual system to distinguish small differences in the illumination falling on adjacent receptors. The retinal image is thought of as being stationary, and it is usually assumed that any motion of the eye is a hindrance to acuity.

Second are the "dynamic" theories. Some of these theories are rather elaborately worked out, and many of them consider central as well as peripheral processes. In these theories the assumption is made that motion of the eye, by means of which the receptors are exposed to first one level of illumination and then to another, is an aid to acuity. The physiological nystagmus has been assigned one or more of several definite roles in these "dynamic" theories. One assumed function of the nystagmus is that of providing the basis for a sort of an "averaging mechanism" which rectifies the blurred image and enhances contrast, thus aiding acuity (**6, 12**). Another assumed role is that of providing a rapid change in the level of illumination falling on the various receptors. This situation presumably results in a higher level of neural activity in the case of certain receptor conductor units than would result from a retinal image which was motionless with respect to the receptor mosaic (**3**).

The results of the present experiments indicate the manner in which the physiological nystagmus influences monocular visual acuity for a conventional grating type test object.

The "Static" Theories

The results of these experiments are in harmony with the hypothesis that acuity would be at its best in a completely immobile eye. However, complete support for this hypothesis requires slight extrapolation from the obtained results. Instead (based on these results), only the general statement may be made that relatively small eye movements result in better acuity than do relatively large eye movements. It must be remembered that no direct evidence has been obtained which would indicate that the immobile eye has better acuity than the eye in motion, for in no case in these experiments was the eye completely immobile during the entire time that a test object was exposed to view. Therefore, the possibility remains that a small amount of motion may actually be an aid to acuity. In general, however, it may be stated that the results of these experiments give support to the hypothesis that the motion of the eye during fixation is a hindrance to acuity. While the results lend no support to the central concept in the "static" theories (that acuity is simply a form of intensity discrimination by adjacent receptors), they are compatible with such a concept.

The "Dynamic" Theories

The results of these experiments are not in accord with the hypothetical roles assigned to the physiological nystagmus in these theories. It should be noted that for the most part the proponents of these theories have made use of the data obtained by Adler and Fliegelman (1) which indicate that the rapid tremor of the eye has average extents of a little over 2 min. of angle of rotation. However, the most recent evidence on the extent of the physiological nystagmus, as summarized in a recent review by Lord and Wright (5), indicates that the extent of the rapid tremor is considerably smaller than that reported by Adler and Fliegelman. A first objection, then, to the "dynamic" theories is the fact that the physiological nystagmus is very much smaller than that assumed, or demanded, by these theories.

A second objection may be made on the basis of the results of the present experiments. It has been demonstrated that drifts of the visual axis greater than 20 sec. of angle of rotation—an angle smaller than the diameter of the average foveal receptor (7)—are a hindrance to monocular acuity. It has also been demonstrated that the relatively large changes in the position of the eye as a result of the rapid tremor are a hindrance to monocular acuity. In general, then, it may be stated that these experiments lend no support to the hypothesis that motion of the eye during the presentation of a test object is an aid to acuity. Neither do these results support the concept that acuity is a form of discrimination of changes in the level of illumination on certain receptors brought about as a result of the physiological nystagmus.

Theoretical Limitations of the Present Study

The present experiments may properly be regarded as exploratory. They do not provide a complete answer to the problem of the role of physiological nystagmus in visual function. Instead, there are a great number of problems remaining to be investigated.

1. Although these experiments have indicated that the nystagmus does not aid monocular acuity with the conventional grating type of test object, these motions may aid in other visual functions such as the perception of straightness

of lines; the recognition of simple borders; or the detection of "graininess."

2. The acuity of the immobile eye has not been investigated, since the eye is not immobile under normal conditions and *E*'s are reluctant to take the rather drastic steps necessary to obtain an essentially motionless eye. However, use of extremely short exposure times would provide records obtained during little or no motion of the eye. Also, some preliminary investigations have indicated that it may be possible, by means of a suitable optical arrangement, to provide an image on the retina which is motionless with respect to the retina itself. This arrangement (equivalent to immobilizing the eye) would permit the use of relatively long exposure times in obtaining data on an effectively immobile eye.

3. In the present experiments the almost instantaneous presentation of the test object in a homogeneously illuminated field results, by itself, in a very high rate of change in the level of illumination on the receptors involved. Therefore, it is conceivable that the eye tremor could still be important in providing a high rate of change in the level of illumination on the receptors during prolonged periods of observation, especially when test objects such as a single fine line are used.

4. Finally, the present results are limited to monocular fixation. Motions of the two eyes during binocular fixation have been recorded in recent experiments (**9**). One very important fact was brought to light in these experiments. *The minute rapid tremor of the two eyes is not coordinated in binocular fixation.* This, of course, means that "corresponding points" on the two retinas are not actually corresponding points. The correspondence must be between *average positions* on the two retinas. It is difficult to see how some of the remarkably low thresholds of only a second or two of visual angle (**2**) are reached when the two eyes are independently undergoing rapid motions of 15 or 20 sec. of arc. The first speculation might be that binocular vision makes use of certain dynamic mechanisms which are not involved in monocular vision.

Summary

Three *S*s fixated monocularly upon crosshairs in a homogeneous field. For a period of 75 msec. this field was replaced by a parallel bar acuity test object after which the fixation field was restored. The test object was presented randomly in horizontal and vertical positions, and *S*'s task was to report the orientation of the parallel bars. At the time of exposure of the test object, the horizontal eye movements were recorded.

The conclusions are as follows: (a) Involuntary drifts of the visual axis during the presentation of the test object are clearly detrimental to monocular acuity; (b) Relatively large amounts of the minute rapid tremor during the presentation of a test object are a hindrance to monocular acuity.

These findings are compatible with the assumption that acuity, as measured by a typical grating test object, depends upon simultaneous differential responses of adjacent receptors. They are not in agreement with the assumption that the small involuntary eye movements contribute to visual acuity by providing temporal variations in the stimulation of each receptor as it is exposed to various parts of the retinal image.

(Manuscript received July 12, 1951)

References

1. Adler, F. H., & Fliegelman, F. Influence of fixation on the visual acuity. *Arch. Ophthal.*, 1934, **12**, 475–483.
2. Berry, R. N. Quantitative relations among vernier, real depth and stereoscopic depth acuities. *J. exp. Psychol.*, 1948, **38**, 708–721.
3. Jones, L. A., & Higgins, G. C. Photographic granularity and graininess. III. Some characteristics of the visual system of importance in the evaluation of graininess and granularity. *J. opt. Soc. Amer.*, 1947, **37**, 217–263.

4. Jones, L. A., & Higgins, G. C. Photographic granularity and graininess. IV. Visual acuity thresholds; dynamic versus static assumptions. *J. opt. Soc. Amer.*, 1948, **38**, 398–405.
5. Lord, M. P., & Wright, W. D. The investigation of eye movements. *Reports on Progress in Physics*, 1950, **13**, 1–23.
6. Marshall, W. H., & Talbot, S. A. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In J. Cattell (Ed.), *Biological Symposia*, Vol. VII, Visual Mechanisms. Lancaster: Cattell Press, 1942.
7. Polyak, S. L. *The retina.* Chicago: Univ. Chicago Press, 1941.
8. Ratliff, F., & Riggs, L. A. Involuntary motions of the eye during monocular fixation. *J. exp. Psychol.*, 1950, **40**, 687–701.
9. Riggs, L. A., & Ratliff, F. Visual acuity and the normal tremor of the eyes. *Science*, 1951, **114**, 17–18.
10. Senders, V. L. The physiological basis of visual acuity. *Psychol. Bull.*, 1948, **45**, 465–490.
11. Walls, G. L. Factors in human visual resolution. *J. opt. Soc. Amer.*, 1943, **33**, 487–505.
12. Weymouth, F. W., Anderson, E. E., & Averill, H. L. Retinal mean local sign; a new view of the relation of the retinal mosaic to visual perception. *Amer. J. Physiol.*, 1923, **63**, 410–411.

EXPERIENTIAL FACTORS IN VISUAL FORM PERCEPTION: II. LATENCY AS A FUNCTION OF REPETITION [1]

RICHARD C. ATKINSON
Indiana University

AND

ROBERT B. AMMONS
University of Louisville

In the first paper in this series (1), a number of meanings of the word "perception" were listed. The discussion was then limited to one meaning, that of visual recognition, and empirical findings were summarized. The present paper summarizes the results of the first of several systematic studies of the latency and probability of occurrence of visual recognition as affected by various experimental manipulations.

It is perhaps in order to indicate the general nature of the phenomena which we propose to study. The term *visual recognition* refers to the kind of experience one has when one becomes aware of a figure. This happens, for example, when one "sees" an ink blot as a familiar object, or sees a reversible perspective in one clear orientation for a moment. A *percept* is a particular familiar organization of stimulus material as figure. An example would be seeing the blot on Rorschach's Plate I (12) as a "bat." The bat would be a percept. More precise, operational definitions will be given these terms later in the paper. In brief, *E* suggests a particular percept, exposes an ink blot, and measures the time necessary to "see" the percept on repeated occasions. To study "learning to learn," *E* observes *S*'s performance with a series of blots and percepts. Finally, by successively introducing two percepts for the same blot, he can study interference.

Several investigators have found a decrease in latency of visual recognition with practice. Leeper's *S*s (9), on the average, required less time to recognize incomplete figures the second time these were presented than they had required the first time. Using a task calling for discovery of a familiar simple figure within a complex figure, Gottschaldt (6), Hanawalt (8), and DuBois and Bunch (2) all observed a decrease in the time to respond, with practice. From this information, and assuming that visual recognition is a learned response, it has been predicted (1) that: (*a*) *With increasing number of exposures, recognition will become more rapid, latency showing a decrease.*

The finding that more complex material takes longer to learn is very nearly universal. McGeoch (10) summarizes a number of studies showing this for such materials as prose, nonsense syllables, finger mazes, and rational learning problems. If we assume that recognition is similar and related to these other processes, and that trials to learn and time taken on a given trial are positively correlated, we can predict that: (*b*) *The more difficult a percept, the greater the latency of recognition.*

There is some evidence for negative transfer from one recognition situation to another. Gibson (5) found that *S*s who learned a second series of nonsense words in the "rest" period between two practice periods with a first list, recognized fewer of the words in the original list on second presentation than did *S*s who had not learned the second list. McKinney (11) found the same sort of thing with pictorial advertisements and nonsense syllables. Here again since we have some indirect evidence and seem to be dealing

[1] The authors gratefully acknowledge many helpful criticisms and suggestions by Dr. Harold Schlosberg, Mrs. C. H. Ammons, and Dr. A. J. Latham.

with learning, we predict by analogy that: (*c*) *The latency of recognition will be greater for a second percept than for the first, when both are based on the same stimulus material.*

People "learn to learn." For example, Ward (13) with nonsense syllables, and Greenberg and Underwood (7) with paired adjectives, reported that *S*s took less time to reach a given criterion of learning, the more lists they had already learned. However, as practice was continued, there was less and less gain. It can be surmised that with sufficient practice a point would be reached where no further list-to-list improvement would appear. Since recognition responses have been made in great numbers by any normal *S*, it is predicted that: (*d*) *With the recognition response, there will be little or no "learning to learn," i.e., little or no decrease in response latency from percept to percept due only to prior recognition experience and not to content.*

Problem

The present study is designed to provide answers to the following questions about the latency of visual recognition responses to Rorschach ink blots: (a) What changes in latency take place as a result of practice? (b) What is the effect of difficulty of perceptual organization on latency? (c) What happens to latency when a second and different perceptual response is being set up to the same stimulus material? (d) Is there a transfer in efficiency with the setting up of successive responses, i.e., a "learning to learn"? In the course of finding answers to these questions, the accuracy of the above predictions can also be checked.

Method

Subjects.—A total of 12 male and 12 female *S*s completed the main experiment. Records of five other *S*s were discarded, for reasons noted later. All *S*s were college students, volunteers not familiar with the Rorschach test, and had taken no advanced psychology courses.

Materials.—A 55- by 72-in. board was mounted vertically on the edge of a table, as a screen. At eye level for a sitting person there was a 7- by 9.5-in. rectangular opening covered by a sliding panel. Behind the panel was a mounting platform for the Rorschach plate. Plates I, IV, V, and VII (12) were used. Responses were timed with a .001-min. Standard Electric timer. In the main experiment, 60 magazine photographs mounted on backing sheets in booklet form served as interpolated material. These photographs were relatively dissimilar to the percepts and ink blots as to content and form.

The apparatus operated as follows: *E* threw a switch completing a solenoid and timer circuit. The solenoid core withdrew, allowing the sliding panel to drop, thus exposing the Rorschach plate at the same time that the timer started. As soon as he "saw" the percept, *S* pressed a button which broke the timer circuit. The timer thus recorded the recognition latency, measured as time from the dropping of the panel by *E* to *S*'s signaling his recognition. The *E* then replaced the sliding panel in front of the Rorschach card.

Procedure.—The *S* seated himself in a chair with his eyes 18 in. from the sliding panel. He was told that he would see some pictures and ink blots, and that he would be told what to "see" in the ink blot, using all parts of it carefully. He was to press a button when he clearly saw what he was supposed to in the ink blot, and not before. The magazine pictures were to be observed between the several exposures of an ink blot, and a single word best describing the content was to be written down.

In the preliminary practice period *S* was instructed to look for a "butterfly," and shown Rorschach's Plate V five times, each time until he indicated recognition. In each period between observations, he looked at one magazine picture and wrote down a descriptive word. After the fifth recognition, the plate was shown again to *S*, and he was asked to point out the major details of the percept, indicating whether or not he had seen the blot this way each time. He was cautioned to use all the blot, the same way each time. Next, *S* was informed that he would see "two women with the backs of their heads together," and Rorschach's Plate VII was exposed inverted. There were five exposure or recognition periods with interpolated magazine picture naming, followed by pointing out of the percept details. The *S* indicated any changes in these details, and was again cautioned to be consistent and complete.

The *S* was then told he would see the "butterfly" plate again, but this time as "two rams bucking each other." Ten exposures were made, separated by 15-sec. inter-exposure periods. Magazine picture naming was continued, and *S* was warned 2 sec. beforehand of the next blot exposure. Once more he was questioned as to blot details and their constancy over the ten exposures, and was cautioned. This completed the pretraining or preliminary practice period.

The regular experimental period followed.

The *S* recognized each of four percepts ten times in sequence. Magazine picture naming was continued in the 15-sec. inter-exposure periods, as was the 2-sec. pre-exposure warning. Rest periods of 2 min. were given between the four percept series. Latency of recognition was recorded from the timer after each of the 40 exposure or recognition periods. After all the recognitions, *S* ranked the four percepts in order of their difficulty for him. He then identified the major details of each percept, and indicated whether they had remained constant. The records of five *S*s who, during the experimental period, did not see the common details, did not see them approximately the same during the ten exposures, or were not able to see some percept at all, were discarded.

A preliminary study of percept difficulty was carried out with a class of 50 students. Each person listed as many whole responses as he could to each of the ten Rorschach plates. To choose plates and percepts for the experiments, frequencies of similar responses were tabulated. After elimination of the colored plates, a search was made for three plates, each with both a "hard" (seldom seen) and an "easy" (often seen) percept, and a fourth plate with a "medium" percept.

For the preliminary practice period, the following plates and percepts were chosen: "butterfly," Plate V (easy, seen by 32 students); "two women with backs of heads touching," Plate VII inverted (medium, seen by 17); "two rams bucking each other," Plate V (hard, seen by 2). Thus practice could be given with easy and hard percepts, and with two percepts based on the same plate.

For the experimental period, the following plates and percepts were chosen: "cat's face," Plate I (easy, seen by 36); "motorcycle," Plate I (hard, seen by 1); "animal skin," Plate IV (easy, seen by 34); and "man seated on a fire hydrant," Plate IV (hard, seen by 2). Thus, there was a hard and an easy percept for each of two plates.

The three practice percepts were always given in the same order. The four experimental percepts were given in completely counterbalanced order, each of the 24 *S*s being randomly assigned one of the 24 possible sequences. Each of the two plates used in the experimental period was exposed 20 times, ten times for each of two percepts. The magazine picture naming and staircase reversal were interpolated between exposures to "break up" visual images and set to remember the percept. The inter-exposure interval was set at 15 sec. because introspective reports from preliminary *S*s indicated that this did not lead to confusion or inhibition.

Results

The data can be analyzed in several different ways. Figure 1 shows the mean and median latencies of recognition for each of the four experimental percepts, as a function of amount of practice (number of exposures). As predicted, (a), latency drops rapidly for the first few trials, then more slowly. Improvement is much more rapid than is the case in the acquisition of many other responses. Means are higher than me-

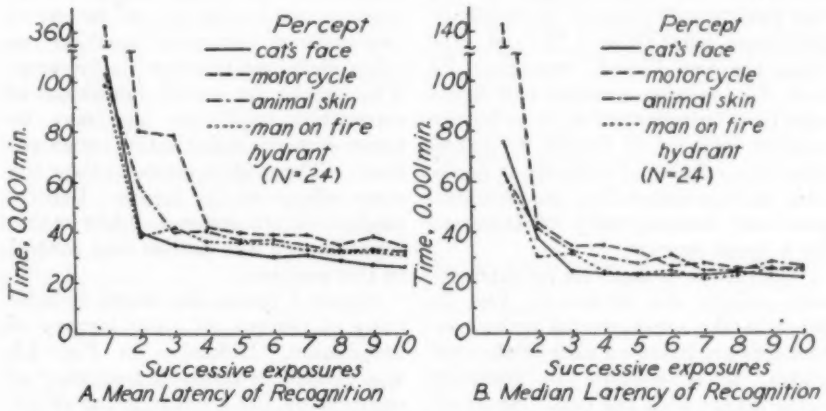

FIG. 1. Mean and median latencies of recognition for each of the four experimental percepts

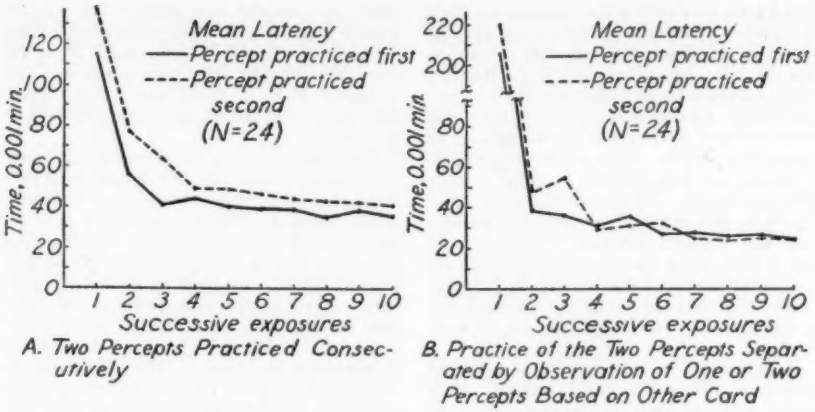
Mean Latency
Percept practiced first
Percept practiced second
(N=24)
Time, 0.00/min.
Successive exposures
A. Two Percepts Practiced Consecutively
B. Practice of the Two Percepts Separated by Observation of One or Two Percepts Based on Other Card

FIG. 2. Interference effects estimated by comparing mean latencies of recognition of percepts presented first with those of percepts presented second from the same card

dians, indicating that the distributions of latencies are skewed. Latencies show about the same sort of change for all four percepts, and relative levels are fairly consistently maintained. Stability of the measures is further indicated by odd-even score reliabilities (product-moment correlations) ranging from .57 for "motorcycle" to .91 for "animal skin," with a median of .76. Corrected to full 10-trial length (4, p. 319), this median *r* becomes .86. Interestingly enough, the two "easy" percepts have higher odd-even reliabilities (.79 and .91) than the two "hard" percepts (.73 and .57), judging easiness and hardness from frequency of mention by the student group. It should be noted that this reliability estimate is probably slightly inflated by interference-produced heterogeneity of latencies for a single percept.

The results of analyses for interference effects are shown in Fig. 2. Since in the experimental series two percepts are based on each of the two plates, it is possible to compare performance with the two. Separate computations can be made for cases where the two percepts based on one plate are given one immediately after the other (24 instances) and where they are separated by one or two percepts based on the other plate (16 and 8 instances, respectively). It can be seen from Fig. 2A, that, as predicted, (c), there is definite interference persisting throughout the practice period, when the second percept based on a given plate immediately follows the first. Figure 2B shows that this interference does not appear, contrary to prediction, (c), when one or two different percepts, based on the other plate, are practiced in between. This might be called inhibition of retroactive inhibition, and may be some sort of disinhibition phenomenon. The median latencies show the same effects as the means. Central tendencies are somewhat less stable because each *S* is not his own control in this analysis.

Figure 3 shows the effects of difficulty of percept on mean latency of recognition. Difficulty in Fig. 3A was defined in terms of frequency of mention by the student group of 50, while difficulty in Fig. 3B was defined

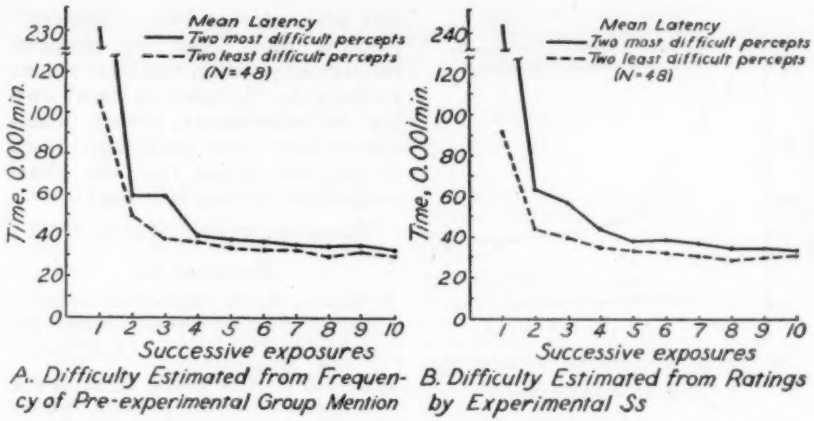
Mean Latency
Two most difficult percepts
Two least difficult percepts
(N=48)
Time, 0.001/min.
Successive exposures
A. Difficulty Estimated from Frequency of Pre-experimental Group Mention
B. Difficulty Estimated from Ratings by Experimental Ss

FIG. 3. Effects of difficulty of percept on mean latency of recognition

as average difficulty ranking by the experimental *S*s. In each case, latencies of both easy and hard percepts show the same sorts of decreases, and, as predicted, (b), the hard percepts have consistently greater latencies throughout the practice. The greatest difference appears at the start of practice. The total time to respond on the first three exposures was correlated with the total on the last seven for each percept. Correlations ranged from .36 for "motorcycle" to .85 for "animal skin," with a median of .58 for all percepts. It can be seen that a percept difficult for *S* at the start of practice tends to remain difficult for him.

As predicted, (d), there was no evidence for a "learning to learn" effect, since mean performance did not improve from the first to the last percept. However, there may have been some improvement during the preliminary practice period when latencies were not recorded. No gross sex differences were observed in latency of recognition.[2]

DISCUSSION

Woodworth (14, pp. 303–305) tells of early experiments similar to this one, with the "*d*-response," measured as time to signal that a familiar picture or object had been recognized. Study of the *d*-response was gradually given up, due mainly to *S*s' inability successfully to discriminate between just seeing "something" and actually recognizing it. Our *S*s reported that they "slipped" a few times, usually toward the end of a series of exposures of a given ink blot. However, this did not happen often, and was usually compensated for by increased caution on the next exposure.

To check on *S*s' reports, a correlation was run for each percept, between total latency on the first three exposures and that on the last seven, as reported in the results section. The median *r* for the four percepts was .58, which indicates that some of the same factors were operating early and late in practice. Mean latencies for the 12 *S*s with the largest total latencies on each percept are compared with those for the 12 with the smallest total latencies in Fig. 4. Curves are of the same shape, flattening out for both groups in the case of every percept. It is hard to believe that the latencies of the slower *S*s in-

[2] In a check experiment involving six *S*s, the interpolated magazine pictures were replaced by an 8- by 10-in. drawing of a reversible perspective, the Schroeder staircase (14, p. 628), and with *S* instructed to speed up the reversals of perspective but with all other conditions the same as in the main experiment. Mean latencies of recognition for this group were about the same as those for the main group, indicating that the type of interpolated material probably does not have much effect on acquisition of the recognition response.

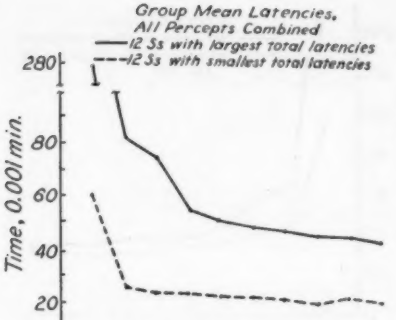

FIG. 4. Mean latencies for slow and fast recognition groups for all percepts combined

cluded very many simple reaction times. The present experiment appears to differ markedly from those with the *d*-response, at the least with respect to nature of material, number of trials, de-emphasis of speed, and introduction of interpolated material between exposures of test stimuli.

It is apparent that set is an important determiner of response in this situation. In a way, it can be said to mediate organizational activity and redintegration. In a general way, the phenomena can be attributed at least partly to the acquisition of specific sets as a function of practice.

SUMMARY

Twenty-four *Ss* practiced visual form perception with two Rorschach ink blots, then were given ten consecutive trials (exposures) with each of four percepts, two based on each of two additional ink blots. Each trial lasted from the time a blot was presented until *S* signalled that he had "seen" the percept suggested by *E* at the start of the particular series of ten trials. Mean and median latencies showed a progressive decrease for each percept, as a function of number of trials, with the greatest decrease taking place on the first three trials. Mean latencies for the second percept based on a blot were greater than those for the first, if there had been no interpolated practice with another blot. "Difficult" percepts took longer to recognize throughout practice, and there was no evidence for "learning to learn" during the experimental period. Findings support four predictions based on the assumption that the visual recognition response is learned.

(Manuscript received May 15, 1951)

REFERENCES

1. AMMONS, R. B. Experiential factors in visual form perception: I. Review and formulation of problems. *J. genet. Psychol.* (In press)
2. DUBOIS, P. H., & BUNCH, M. E. A new technique for studying group learning. *Amer. J. Psychol.*, 1949, **62**, 272–278.
3. ELLIS, W. D. (Ed.) *A source book of Gestalt psychology.* New York: Humanities Press, 1950.
4. GARRETT, H. E. *Statistics in psychology and education.* (2nd Ed.) New York: Longmans, Green, 1938.
5. GIBSON, J. J. Retroaction and the method of recognition. *J. gen. Psychol.*, 1934, **10**, 234–236.
6. GOTTSCHALDT, K. Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren: I. Über den Einfluss gehäufter Einprägung von Figuren auf ihre Sichtbarkeit in umfassenden Konfigurationen. *Psychol. Forsch.*, 1926, **8**, 261–317. (See Ellis [3] for an extensive English summary.)
7. GREENBERG, R., & UNDERWOOD, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, **40**, 452–457.
8. HANAWALT, N. G. The effect of practice upon the perception of simple designs masked by more complex designs. *J. exp. Psychol.*, 1942, **31**, 134–148.
9. LEEPER, R. A study of a neglected portion of the field of learning—the development of sensory organization. *J. genet. Psychol.*, 1935, **46**, 41–75.
10. MCGEOCH, J. A. *The psychology of human learning: an introduction.* New York: Longmans, Green, 1942.
11. MCKINNEY, F. Retroactive inhibition and recognition memory. *J. exp. Psychol.*, 1935, **18**, 585–598.
12. RORSCHACH, H. *Psychodiagnostics: Plates.* (4th Ed.) Berne: Hans Huber.
13. WARD, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 4, Whole No. 220.
14. WOODWORTH, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

THE EFFECT OF SIMULTANEOUS VISUAL STIMULATION ON ABSOLUTE AUDITORY SENSITIVITY[1]

LEE W. GREGG AND W. J. BROGDEN

University of Wisconsin

A survey of the literature dealing with the general problem of intersensory relationships reveals that significant changes in the sensitivity to a primary stimulus have been reported when an auxiliary stimulus of supraliminal intensity is presented simultaneously to receptors of a sense modality different from that to which the primary stimulus was presented. However, the changes have not always been found to lie in the same direction. In some instances, the auxiliary stimulus has been found to enhance or facilitate sensitivity. In others, an inhibitory effect has been observed. Studies in this general area of research have been reviewed by Ryan (9) and by Gilbert (6).

The present experiment is an attempt to differentiate conditions under which an auxiliary light stimulus facilitates or inhibits sensitivity to auditory stimuli. There are relatively few studies that are directly relevant to this problem. Hartmann (8) measured pitch and loudness discrimination by means of the Seashore records under conditions of very high and very low levels of general illumination. He reported a slight increase in acuity for the high level of illumination. However, in another experiment in which *S*s were tested in complete darkness and were required to make a verbal report instead of recording their responses in the usual way, auditory acuity was greater than for the high level of illumination. Hartmann proposed that "a simultaneous auxiliary heteromodal stimulus augments the main stimulus when the former is in the background, but it acts in an inhibiting way when it is focal." Cason (3) reported a facilitating effect when the onset of an auxiliary stimulus coincided with the onset of the main stimulus. He had 40 *S*s judge the intensity of a second stimulus with respect to a first, and found that when a light was presented, the auditory stimulus was judged louder than when the same auditory stimulus was presented without light. Similarly he found that a sound presented with a light made the light appear brighter than when the same light was presented alone. Child and Wendt (4) investigated the influence of a light upon auditory acuity as a function of the temporal interval between the light and tone. They found that light facilitated reception of the tone and that the maximum effect was obtained when the light preceded the tone by .5 sec. These authors favored an interpretation of their results in terms of some type of neural summation, but nonetheless did not exclude the possibility of peripheral factors and "complex phenomena of attention and judgment."

Preliminary experiments by the present authors indicated that intersensory effects could be measured by comparing auditory or visual thresholds to tone and light separately, with thresholds for tone and light in combination. In these experiments, separate thresholds were determined by a modification of the method of limits. Then, the two stimuli, a 1000-cycle tone and a circular light patch, were presented simultaneously, and as *E* varied the intensities of the two stimuli, *S* reported on the presence of the light, tone, or both. The group mean thresholds for tone and light were not significantly different for the two conditions. However, examination of the thresholds obtained with the combination of tone and light revealed that for many *S*s one of the two thresholds was lower than that for presentation of the stimulus alone; the other threshold was equal to or higher than that obtained when the stimulus was presented alone. Only in rare cases were both thresholds for the combined stimuli lower than for their separate presentation alone. Facilitation of auditory acuity by light occurred with about the same frequency as did inhibition of auditory sensitivity. A similar relation between facilitation and inhibition of visual sensitivity by the tone was found also. Obviously, the experimental conditions in which tone and light were presented together did not provide for separation of the factor or factors that produce inhibition or facilitation. Either light or tone may have either function. It is a reasonable assumption that the instructions to report on both the tone and the light, and the information that both

[1] Supported in part by the Research Committee of the Graduate School from special funds voted by the State Legislature.

auditory and visual thresholds would be measured, include the factors that produce facilitation or inhibition. One can assume further that a facilitating effect would be more likely to occur if the *Ss* are oriented toward one of the two stimuli only, and if thresholds for this stimulus only are measured, and *S* is so informed. It follows, then, that an inhibitory effect should have greater likelihood of occurrence if *S* is oriented toward both kinds of stimuli, even though thresholds are obtained for one only, and *S* is so informed. It is these hypotheses that determined the design of the present experiment. Thresholds of auditory acuity were obtained for two groups of *Ss*. One group was informed of the light condition that might accompany the tone and was required to respond to the light as well as to the tone. The other group was not informed about the light and no request to report on it was made. For both groups, auditory thresholds were obtained under a no-light condition and under two light conditions that were simultaneous with the presentation of each tone and that differed from each other only in terms of brightness. The brightness of the light in both cases was near the threshold.

Method

Apparatus.—The *S* sat in a sound-treated room at a table facing one wall. A chin rest was attached to the table in such a position that *S*'s head was level with a gray metal box, 11 by 17 in., mounted on the wall and containing the light source. A pair of loudspeakers was mounted beneath the table, one at each end so that each speaker was approximately the same distance from *S*'s head. A ceiling fixture and globe containing a 100-w. bulb provided the illumination for the room. A microphone rested on the table to the left of *S*, out of his line of vision. A large loudspeaker for intercommunication was affixed to the wall at *S*'s right. The rest of the apparatus was located in an adjoining room occupied by *E*.

The auditory stimulus, a 1000-cycle tone, was provided by a Wien bridge oscillator carefully calibrated by means of a standard frequency broadcast by the National Bureau of Standards' radio station WWV. The output of the oscillator was fed into a Hewlet-Packard attenuator set, model 350A, then to a matching transformer and into the loudspeakers mounted beneath *S*'s table.

The light stimulus was an increase in the illumination of a small circle ¼ in. in diameter in the center of the metal box on the wall facing *S*. The patch subtended approximately 43-minutes visual angle. An opal glass plate was mounted behind the ¼-in. circle cut in the box, and behind the glass plate was a neutral tint filter, No. 2.0. A 6-in., 4-w. fluorescent bulb (G. E., daylight) was mounted in the box and activated by a DC voltage source. A fluorescent bulb was used so that the intensity of the light source could be changed by varying the voltage without the corresponding hue change observed with incandescent bulbs under these conditions. The fluorescent bulb was operated at the lowest current possible without having the tube cut off. The light stimulus was therefore an increase in brightness of the circle brought about by shorting out fixed resistors in series with the source. The brightness of the circle, which is referred to as the fixation circle in this experiment, at the lowest current was 2.050 ml. as measured by a Macbeth illuminometer.

Experimental design.—Three light conditions were used. These were a zero light condition with the fixation circle consistently at a brightness level of 2.050 ml., an increase in brightness of .015 ml., and an increase in brightness of .055 ml. The latter two values were chosen on the basis of the data obtained from the 54 *Ss* used in the preliminary studies. Differential brightness thresholds were obtained on these *Ss* by the same technique used to determine the auditory thresholds in this experiment, a modification of the method of limits (**5**). Figure 1 presents the distribution of these thresholds expressed as $\Delta I/I$. The ordinate is per cent frequency of occurrence of the thresholds. The brightness change of .015 ml. corresponds to log $\Delta I/I$ of −2.14, and from Fig. 1 it is seen that approximately 38% of the thresholds fall below this

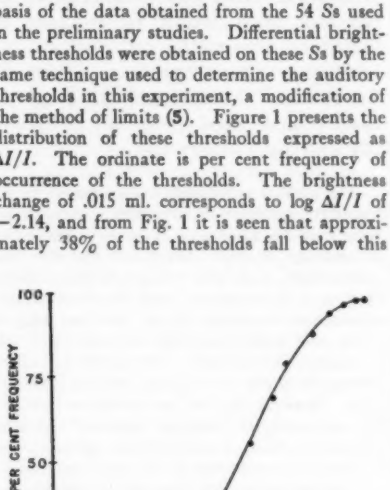

FIG. 1. Distribution of 676 difference thresholds to light for 54 *Ss*. The value of *I* is 2.050 ml. The solid curve was fitted by eye to the empirical points.

value. The increase of .055 ml. corresponds to the log $\Delta I/I$ value of -1.57; approximately 100% of the cases falling below this value.

The minimal audible threshold for the 1000-cycle tone was measured by a modified form of the method of limits (5). Steps of 2 db between stimulus intensities were used. The threshold intensity of the tone for 50% response was recorded in terms of db of attenuation. The raw data consisted of threshold values expressed in terms of the number of db of attentuation from a base tone which was measured at 55 db above the reference level of .0002 dynes/cm.^{2} by means of a General Radio sound level meter Type 759, weighting position *A*. The thresholds were corrected to the reference level by subtracting the raw thresholds from 55 db. A warning buzzer whose duration was 300 msec. preceded the tone (alone or with a light increase) by 1000 msec. A set of electronic timers permitted this temporal pattern to be rigidly followed. The duration of the tone was 2000 msec.

Subjects and procedure.—The *S*s, 18 men and 18 women, were students in introductory psychology at the University of Wisconsin. For each *S*, nine thresholds for the tone were determined, three for each level of the auxiliary light stimulus. The six permutations of the three levels gave the six sequences to each of which three *S*s were randomly assigned. Thus, 18 *S*s made up Group I operating under instructions to report the presence of the light as well as that of the tone. Another 18 *S*s made up Group II, and they were instructed only to respond to the tone. Upon reporting for an experiment on auditory acuity, *S* was conducted to the experimental room. He was asked to be seated at the table and the following instructions were read:

> This is an experiment on auditory acuity as you know. First I want you to listen to a series of faint tones. Please place your chin on the rest and keep it there throughout the experiment. It is necessary to hold constant the position of your head with reference to the loudspeakers. The buzzer is a warning that a tone will follow. When the buzzer sounds, fixate the small gray circle with your eyes. This is also necessary to help you maintain a fixed orientation with respect to the loudspeakers. As soon as you hear the tone, let me know by saying "tone."

Five practice tones approximately 18 db above threshold were then given. If *S* was assigned to Group II, no mention was made of the light stimulus. The *E* simply left the experimental room and the threshold determinations were begun. If *S* was assigned to Group I, the following paragraph was read also:

> To be sure you are maintaining the fixed orientation, the fixation circle will sometimes be lighted simultaneously with the sound of the tone. If you see the light, report it after you have reported the occurrence or nonoccurrence of the tone.

Then *E* left the room and the thresholds were measured. After the nine thresholds to the tone were obtained, *E* gave the following additional instructions to *S* after first establishing whether or not the light had been noticed:

> I have taken all the auditory thresholds needed for the experiment, but I would like to get some additional data on the fixation light. When the buzzer sounds, look at the fixation circle and let me know whether or not you see it light up. Just say "light" or "no light" on each trial.

Then *E* presented five zero light trials, five brightness increases of .015 ml., and five brightness increases of .055 ml. in a random order to *S* and these responses were recorded.

Results

For each *S* there were nine threshold measures, three each for the zero, .015 ml., and .055 ml. brightness increases. The median threshold for the zero light condition was taken as the zero sensation level for that *S*, i.e., his normal threshold of audibility. From this value each of the nine individual thresholds was subtracted to give the raw data in terms of sensation level.

Separate analyses.—For purposes of easier computation, an arbitrary constant, 6, was added to each threshold expressed in terms of sensation level to eliminate negative numbers. These data were entered into a multiple classification table, and separate analyses were made for each group as well as for the combined data. Table 1 presents the results of the separate analyses for the two groups (10). The main effect concerned with the variation in brightness increases is referred to as light conditions in the table; sequence refers to the ways in which the light conditions were or-

TABLE 1

SUMMARY OF RESULTS OF ANALYSES OF VARIANCE

Source	*df*	Group I		Group II	
		ms	*F*	*ms*	*F*
Light conditions	2	8.50	6.16**	11.00	11.22**
Between no-light and two levels of light	1	15.00	10.87**	15.00	15.31**
Between levels of light	1	2.00	1.45	7.00	7.14*
Sequence	5	11.06	1.11	22.00	3.88*
Trials	2	6.50	4.71*	9.00	9.18**
L × S	10	13.70	2.28*	5.20	1.89
L × T	4	2.75	1.99	1.50	1.53
S × T	10	2.70	1.96	2.00	2.04*
L × S × T	20	2.15	1.56	1.01	1.03
Ss within sequences	12	9.92	7.19**	5.67	5.79**
Ss × L	24	6.00	4.35**	2.75	2.81**
Ss × T	24	0.63	0.46	1.04	1.06
Ss × L × T	48	1.38	—	0.98	—
Total	161				

* Significant at the 5% level of confidence.
** Significant at the 1% level of confidence.

dered; and trials refers to the first, second, or third thresholds for a given light condition. The interactions are designated by the first letter of the words describing the main variables, e.g., L × S designates the interaction of light conditions with sequence.

For Group I, light conditions is significant above the 1% level of confidence. The mean deviations of the auditory thresholds from the zero sensation level for this group are 0.00, .78, and .52 db for the zero, .015, and .055 ml. brightness increases respectively. The 2 *df* for the light conditions were broken down into a between no-light and the two levels of brightness increases with 1 *df*, and a between levels of brightness increases with 1 *df*. The comparison in the former case is one between the mean threshold for the zero light condition and the mean of the mean thresholds for the two levels of brightness increases, and is significant above the 1% level of confidence. However, the comparison between levels does not yield a significant *F*. The effect of the instructions to this group, mention of the light and requiring a report of it, is to raise the auditory threshold, but there is no significant differential effect on sensitivity for the two brightness levels.

For Group II, light conditions is also significant above the 1% level. But examination of the mean deviations from the zero sensation level shows a completely different trend in the direction of change from that for Group I. The comparison of the group means is most clearly shown in Fig. 2. The solid line connecting the filled circles shows the relation between mean deviations and light conditions for Group I. The broken line connecting the unfilled circles is for Group II. The mean deviations for the latter are −.06, −.46, and −.96 db for the zero, .015, and .055 ml. brightness increases respectively. When the sums of squares for light conditions is broken down into those for the individual *df*, the difference between the no-light and two levels of brightness increases is significant above the 1% level, and the difference between levels of brightness increases is significant above the 5% level. For

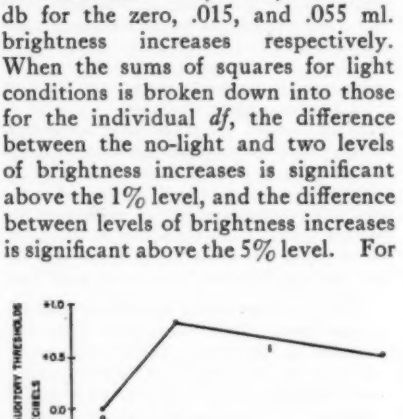

FIG. 2. Group mean auditory acuity is plotted as a function of the magnitude of the brightness change of the auxiliary visual stimulus. The measures of auditory acuity are deviations from the group mean of the median sensation level of each *S* for the auditory thresholds at the zero brightness change of the light stimulus.

Group II, the effect of the auxiliary light stimulus has been to lower the thresholds, with a differential effect observed between the levels of brightness increases.

For both Group I and Group II, the *F* for trials is significant. This indicates a practice effect, a decrease in the magnitude of the thresholds as successive thresholds were obtained. The effect is general since the interactions of trials by *Ss* within sequences is not significant. Sequence in this design is not orthogonal to *Ss*. To test the significance of this variable, the mean square for sequence must be divided by the mean square for *Ss* within sequences since the latter source of variation is significant. For Group I, sequence is not significant; for Group II, it is. The significant *Ss* within sequences variation, above the 1% level for both groups, reflects the usual occurrence of individual differences. However, not only was this source of variation present, but also different *Ss* reacted differently to the light conditions. The *Ss* within sequences by light conditions interactions for both groups is significant above the 1% level.

Two other interactions were significant or very nearly so for both groups. These were the L × S and S × T. They reflect differential variation in the light conditions and trials as a function of sequence, a type of practice effect. However, the use of the complete set of six sequences counterbalances this effect for the light condition means.

Combined analysis.—Although the combined analysis presented in Table 2 provides essentially the same results as those obtained from the separate analyses, it does permit the evaluation of the over-all effect of instructions on the group means. Since different *Ss* appear in the two groups, i.e., since

TABLE 2

SUMMARY OF RESULTS OF ANALYSIS OF VARIANCE FOR COMBINED DATA FROM GROUPS I AND II

Source	*df*	ms	*F*
Groups (instructions)	1	70.00	8.99**
Light conditions	2	4.00	3.39*
Sequence	5	30.80	3.95**
Trials	2	14.50	12.29**
G × L	2	15.50	3.54*
G × S	5	2.80	0.36
G × T	2	1.00	0.85
L × S	10	15.40	3.52**
L × T	4	1.50	1.27
S × T	10	2.40	2.03*
G × L × S	10	3.50	0.45
G × L × T	4	2.75	0.63
G × S × T	10	2.30	0.30
L × S × T	20	1.25	0.29
G × L × S × T	20	2.00	1.69
Ss within sequences	24	7.79	6.60**
Ss × L	48	4.38	3.71**
Ss × T	48	0.83	0.70
Ss × L × T	96	1.18	—
Total	323		

* Significant at the 5% level of confidence.
** Significant at the 1% level of confidence.

Ss are not orthogonal to groups (instructions), we must use the mean square for *Ss* within sequences to test the effect of instructions. This *F* is significant at the 5% level of confidence. The *F* for the light conditions is significant at the 5% level, but it should be pointed out that in this analysis this effect represents the differences in the means for the light conditions summed over both groups. It is the interaction of groups by light conditions that describes the difference with which we are primarily concerned. The *F* for this interaction is significant at the 5% level. This indicates a differential effect from one light condition to another as we move from one set of instructions to another. In this analysis, sequence, trials, the L × S interaction, and the S × T interaction previously observed

to be significant in the separate analyses are found to be significant.

In the type of analysis carried out for these data, the light treatments, sequence, trials, and instructions were treated as separate factors. An analysis involving the latin square principle also is possible since the six sequences of the light conditions represent all permutations of these conditions. Thus, each light condition occurs once in each row and twice in each column. This type of analysis, described by Grant (7), was carried out for the separate groups, and the results are essentially the same as those already stated for the main effects. However, an analysis of this type permits the evaluation of an ordinal position effect. If we sum the threshold measures over trials and light conditions for the first, second, and third position of the sequence, we find for Group I the following mean values: +1.37, +.43, and −.50 db; and for Group II we find these means: +.04, −.39, and −1.13 db. The effect in both cases is one of lowering the auditory thresholds as ordinal position in the sequence increases. The *F* values for Group I (47.50/2.52 = 18.85) and for Group II (18.50/1.46 = 12.67) are both significant at the 5% level with 2/138 *df*. This finding corroborates the indication of a practice effect brought out by the significant effect of trials, and shows that not only is the effect present within a given light condition, but that it also extends over the whole series of threshold measures.

The series of trials in which the light conditions were presented alone at the end of the threshold measures was obtained in order to evaluate roughly the degree to which it was an effective stimulus. As pointed out in the procedure section the brightness values of the light stimulus were chosen on the basis of data obtained on a group of *S*s similar to those used in the present experiment. It was expected that effectiveness of the brightness increases would vary from individual to individual, but inasmuch as a relatively large group of *S*s was used, we would suppose that on the average, responses to the light would be comparable from experiment to experiment. For the zero, .015, and .055 ml. brightness increases we found the following percentages of reports: for Group I, 3.3, 36.6, and 87.7% respectively; for Group II, 15.5, 37.7, and 72.2% respectively. No attempt was made to analyze these results. Probably the most interesting feature of these results is the number of reports of brightness increase for the condition in which there was no change of the light. A possible rationalization is that by the end of the experimental session a very strong response tendency was established and set off by the warning buzzer that preceded each stimulus presentation. This may represent another variety of sensory conditioning (1, 2).

Discussion

The results of the present experiment clearly substantiate the hypothesis that an auxiliary contiguous light stimulus decreases sensitivity to a 1000-cycle tone when *S* is instructed to report the presence of the light as well as the tone, and increases sensitivity to a tone when *S* is not required to report on the presence of the light. In the latter case, the degree of facilitation of auditory acuity increased as the brightness change of the light was increased from .015 to .055 ml. The magnitude of the decrement in auditory acuity was not found to be different for the two conditions of brightness change.

Although there is a clear distinction between the two sets of instructions that produced facilitation on the one hand and inhibition on the other, it is difficult to give a more precise statement of the variable or variables concerned, particularly with respect to how these conditions were maintained for the duration of the experimental session. No gain in precision is made by

stating that the two kinds of instructions produced two different "sets," one of which is responsible for facilitation and the other for inhibition. Hartmann's (8) proposition that an auxiliary heteromodal stimulus when focal inhibits sensitivity to the main stimulus and when marginal facilitates the main stimulus is a general statement that may describe the difference between our two kinds of instructions. It adds no more to our knowledge of these kinds of intersensory effects than does use of the concept of set.

Obviously, the instructions produce two different classes of behavior in the *Ss*, behavior that persists for the duration of the experimental session, and in one case facilitates auditory acuity and in the other case inhibits auditory acuity. Our only measurement of the difference in behavior is in terms of the *Ss*' report of the light and the frequency with which they report the tone at the different db levels. Both groups of *Ss* oriented to the fixation circle at the sound of the warning buzzer, and it is likely that, because of the different instructions, some response evoked by the buzzer is differential for the two groups. Our experimental procedure did not provide for direct observation of the *Ss* during the experiment, and discussion with the *Ss* at the close of the experiment did not indicate that they were aware of any such responses. Therefore, we have no clue as to the nature of such responses, assuming that they did occur. One could speculate in terms of responses that provide for orientation to visual and auditory stimuli, or for increments or decrements in general bodily tonus. It is also possible that responses primarily neural in nature are involved and that interactions of these responses within the nervous system provide the basis for distinguishing the phenomena of facilitation and inhibition of auditory acuity.

Perhaps a more fruitful approach toward a more precise definition of the conditions of the phenomena will be obtained by considering the function of the magnitude of the brightness change of the auxiliary light stimulus. The two degrees of brightness change used in the present experiment were small and were in the neighborhood of the magnitude of the DL. Further experimentation should be directed towards increasing both the number and the range of brightness values for the accompanying light stimulus. It is reasonable to expect that the magnitude of the facilitation of auditory acuity will be affected by extension of this variable, and it is possible that the magnitude of the inhibitory effect might be affected also. With marked increases in the brightness of the visual stimulus, the differential effect of instructions on auditory acuity might disappear. Since responses to intense stimuli are more likely to occur than are responses to weak stimuli, other things being equal, a progressive decrement in auditory acuity might be obtained as a function of increasing the brightness of the visual stimulus. An important parameter in these hypothetical relationships is the degree of practice by *S* in auditory threshold determination. Our data show a significant effect on auditory acuity both for trials and ordinal position of brightness of the auxiliary visual stimulus. Increasing the amount of practice, that is, the number of auditory thresholds taken for each *S*, might to some extent counteract the effects on auditory acuity that have been hypothesized for increments to the brightness of the visual stimulus. Experiments involving these two variables represent the next step in analyzing the intersensory effects that have been demonstrated in the present experiment.

Summary

Auditory thresholds were measured under three conditions of auxiliary visual stimulation. A small light patch was fixated by *S* and increases in the brightness of the patch of zero, .015, and .055 ml. were used. For each *S* three auditory thresholds were determined for each of these light conditions. One group of 18 *Ss* received instructions to report the presence or absence of the light as well as that of the tone; another group of 18 *Ss* made no specific response to the light and were instructed to report presence or absence of the tone only. The results showed a significant elevation of the auditory thresholds for the light increases when a specific verbal response was required to the auxiliary stimulus and a significant lowering of the thresholds as the brightness of the light increased, when no response to the auxiliary stimulus was required.

(Manuscript received June 4, 1951)

References

1. Brogden, W. J. Sensory pre-conditioning of human subjects. *J. exp. Psychol.*, 1947, **37**, 527–539.
2. Brogden, W. J. Sensory conditioning

measured by the facilitation of auditory acuity. *J. exp. Psychol.*, 1950, **40**, 512–519.

3. Cason, H. Sensory conditioning. *J. exp. Psychol.*, 1936, **19**, 572–591.
4. Child, I. L., & Wendt, G. R. The temporal course of the influence of visual stimulation upon the auditory threshold. *J. exp. Psychol.*, 1938, **23**, 109–127.
5. Culler, E., Finch, G., Girden, E., & Brogden, W. Measurements of acuity by the conditioned response technique. *J. gen. Psychol.*, 1935, **12**, 223–227.
6. Gilbert, G. M. Inter-sensory facilitation and inhibition. *J. gen. Psychol.*, 1941, **24**, 381–407.
7. Grant, D. A. The latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427–442.
8. Hartmann, G. W. The facilitating effect of strong general illumination upon the discrimination of pitch and intensity differences. *J. exp. Psychol.*, 1934, **17**, 813–822.
9. Ryan, T. A. Interrelations of the sensory systems in perception. *Psychol. Bull.*, 1940, **37**, 659–698.
10. Snedecor, G. W. *Statistical methods.* (4th Ed.) Ames: Iowa State College Press, 1946.

THE PERCEPTION OF OBSTACLES BY THE DEAF

PHILIP WORCHEL AND JOE H. BERRY

University of Texas

A series of controlled experiments at the Cornell Psychological Laboratory (**2, 3**) leads to the conclusion that auditory stimulation is both a necessary and sufficient condition for the perception of obstacles by the blind.

Because this conclusion from the deaf-blind (**3**) was based upon the performance of only ten *Ss* whose handicaps and ages (range from 34 to 55 yr., mean 45.9 yr.) make them a rather unique group, it was deemed advisable to extend the experiment. The crucial point in the procedure involved the determination whether the *Ss* could learn to perceive a movable obstacle under laboratory conditions (**3**). Obviously, *Ss* blind from birth or blind for at least 25 yr. who had not learned to perceive obstacles could not be expected to do so under the limited time allowed in the laboratory. Habits interfering with acquisition of this "obstacle-sense" could not be eliminated in a brief time, and thus factors other than the absence of hearing might have been responsible for their inability to learn to perceive obstacles. In addition, the findings should be confirmed under conditions usual to the blind, since an *S* tested indoors might receive cues not present outdoors or be deprived of cues usually present in the blind *S*'s environment.

Thus it is desirable to repeat the crucial learning experiment of Worchel and Dallenbach (**3**) under outdoor conditions with a group of young *Ss* who are not blind. Seeing-deaf persons with normal external meatuses and tympanums meet these conditions. By blindfolding these *Ss* we are able to approximate the conditions of the deaf-blind.

Method

The experiment was performed on a concrete walk 60 ft. long and 6 ft. wide. The walk is situated on a plot of ground which is located approximately 50 ft. from a heavily traveled highway, and 30 ft. higher than the road. The walk lies in a perpendicular direction to the highway. It is bordered on both sides by grass; and large oak trees occupy the surrounding area. Some 200 ft. away is a school building.

A considerable amount of noise was present in the experimental area. From across the street pneumatic drills, electric saws, and other such noises associated with building activity could be heard periodically. Half a block away cars and especially huge moving vans created a tremendous amount of noise stopping and starting at a stop light. The intensity of sound in the area was recorded over a period of eight days by a Type 759 Sound Level Meter (General Radio Company), and was found to range from 32 db during quiet periods to 68 db when moving vans, city busses, cars, etc., were prevalent in the area. The average intensity over the eight-day period was 46.8 db.

The experiment was conducted at two periods of the day: in the morning from 9:00 to 12:00, and in the afternoon from 2:00 to 5:00. As a rule, *Ss* served at both times of the day, though in several cases they were tested only at one period. When *S* served at only one period, the experimental hour was arranged so that he would be tested during the same time each day.

A series of 30 trials was given to each *S* during each experimental period every day except Sunday. These trials were continued until *S* gave 25 reports in any one series without coming into contact with the obstacle. In no case was the number of series extended beyond seven (210 trials).

Apparatus.—In all series the obstacle was a ¼-in. masonite board which was 4 ft. wide and 4 ft. 10 in. high. This board was attached to a frame made of ¾-in. by 1¾-in. oak wood. The frame was mounted on 1¾-in. rollers so that the obstacle could be mobile. When the board was mounted on this frame, its lower edge was 2 ft. above the floor.

The blindfold used was a pair of blackened U. S. Navy Variable Density Goggles. These goggles fit snugly to the face, and possess a knob which protrudes from them. In case of collision,

the knob is long enough to protect the face, and so act as a convenient shock absorber.

Audiograms were obtained on all *Ss* with a Western Electric portable 6BP audiometer. This particular instrument emits frequencies between 100 and 10,000 cy./sec., and is calibrated in the following frequencies: 128, 256, 512, 1024, 2048, 4096, 8192, 9747, and 10,000 cy. Half-octaves and quarter-octaves are also indicated.

Subjects.—Arrangements were made with 8 males and 7 females to serve in the study. Ten with hearing deficiencies were selected from the students at the Texas State School for the Deaf.[1] The other five had normal hearing or better and were students at the University of Texas. The *Ss* had no knowledge of investigations in obstacle perception.

A summary of the physical conditions of the deaf *Ss* is given in Table 1.

The two *Ss* lacking vestibular function could not stand on one foot for 30 sec., and when rotated in a conventional rotation chair, they failed to exhibit the behavior which normally characterizes semicircular canal stimulation.

If the meatuses and tympanums are capable of being stimulated by air and sound waves, they would be stimulated in all these *Ss*.

Procedure.—Each day when *S* reported to the experimental area he was given a copy of instructions to read. After he had thoroughly familiarized himself with what he should do, he was blindfolded, led about the experimental area to disorient him, and then placed at one of the various predetermined starting points.

The edge of the walk was chalked off in feet from zero to 48 ft. The *S* would be placed at points 0, 3, 6, 9, or 12 ft. from the zero mark on the walk; and the obstacle was placed at distances of 6, 12, 18, 24, 30, or 36 ft. from *S*. The starting points and the obstacle distances were both varied from trial to trial. The haphazard order of the 30 trials was the same for all *Ss* during any one day, but varied from day to day. Every position was used as often as every other in any one series.

There are several reasons for using a movable obstacle. Previous experiments, especially the 1944 investigations (2), indicated that so long as *S* approached a fixed obstacle, the unevenness of the floor (or walk) may aid in locating the obstacle. The *S* soon learned the location of the various cracks and irregularities of the path, and from his familiarity with the path he could tell approximately how far he was from the fixed obstacle. With this information, *S* could then estimate the distance through which he walked and stop just short of colliding with the obstacle. By using a movable obstacle, *Ss* could not possibly ascertain the obstacle's position by secondary cues from the walk.

After being thoroughly disoriented, *S* was placed at the designated starting point facing the obstacle. At a tap on the left shoulder, he walked toward the obstacle. When he perceived the object (first perception), he stopped momentarily and raised his right arm. He then lowered his arm and continued walking until he was as close as he thought he could get without touching it (final appraisal). At this point he raised his left arm. The distance from the obstacle to the point where *S* indicated his first perception was measured to the nearest foot. The distance at which *S* reported he was as close to the board as possible was measured to the nearest 3 in. After the completion of a trial, *S* was again walked by *E* in a circuitous route on the grass to disorient him, and once again he was placed on a new starting point. At the same time, an assistant moved the board to a new position. To avoid any constant cues in moving *Ss* and placing them on the starting points, *E* and the assistant traded functions from day to day.

Early in the series of experiments we noticed that if the obstacle happened to be located between *S* and the sun, *S* invariably signaled a perception when he came within the shadow cast by the obstacle. This secondary cue was obvious. When *S* walked into the shadowed area on a hot, sunny day, the change to a cooler temperature was easily detected. To eliminate this cue, he was always placed between the sun and the obstacle.

TABLE 1

PHYSICAL CONDITIONS OF DEAF *Ss*

S	Age (years)	Age at Deafness	Per Cent Hearing Loss	
			Left	Right
DA	21	2	52	52
BC	17	congen.	38	32
EC	18	1–2	60	45
JE*	16	11	100	100
WF	17	congen.	68	48
BG	21	congen.	62	68
MH	19	congen.	100	100
PH*	18	5	100	100
ML	17	congen.	49	49
JM	20	2	40	45

* These *Ss* had no vestibular sensitivity; all others had normal vestibular sensitivity. All *Ss* had normal external ears.

[1] The cooperation of Mr. O. L. McIntyre, principal of this school, in allowing us the facilities of the school and in acting as interpreter between *E* and the deaf-mute *Ss*, is gratefully acknowledged.

TABLE 2

MEAN DISTANCE (IN FEET) OF "FIRST PERCEPTIONS" AND "FINAL APPRAISALS" AND MEAN NUMBER OF COLLISIONS IN EACH SERIES OF 30 TRIALS FOR DEAF (N = 10) AND HEARING (N = 5) *Ss*

Ss	Report	Series													
		1		2		3		4		5		6		7	
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Deaf	First perception (p)	7.26	3.87	7.57	4.27	6.64	3.66	6.08	4.11	7.63	4.14	8.05	3.33	8.64	2.43
	Final appraisal	6.77	3.22	7.12	3.56	5.87	2.96	5.14	2.82	6.33	3.72	6.19	2.87	6.91	3.63
	Collisions														
	Before (p)	20.8	8.48	16.2	9.23	15.8	7.34	14.3	7.81	13.0	7.95	9.9	5.66	10.6	7.58
	After (p)	4.2	3.71	6.8	5.65	5.8	4.17	5.1	3.86	5.5	4.79	6.6	3.97	5.4	4.17
Hearing	First perception (p)	9.57	1.72	9.80	2.66	7.95	.94								
	Final appraisal	3.22	1.67	1.85	.83	.90	.47								
	Collisions														
	Before (p)	0.6	.49	0	—	0	—								
	After (p)	5.0	2.76	1.8	1.32	1.6	1.74								

In order to facilitate learning as much as possible, the following rules of procedure were utilized: (a) *punishment*—*S* was not prevented from colliding with the obstacle; (b) *reward*—if *S*'s appraisals were 3 ft. or less from the obstacle, he was praised by a pat on the back, or *E* would take *S*'s hands and form the manual sign for "good"; (c) *withholding reward*—whenever appraisals were more than 3 ft. or the performance was much poorer than previous ones, praise was omitted; (d) *knowledge of results*—after every perception, *S* was led up to the obstacle so that he could know the amount of his error; (e) *knowledge of task to be completed*—during the latter half of the experimental period, *S* was informed from time to time of the number of trials yet to be made; (f) *avoiding fatigue*—care was taken to insure that *S* was not exhausted during the trials of an experimental hour.

Instructions.—When *S* arrived at the experimental area, a typewritten copy of the instructions was given to him to read and study. These instructions were the same as used in the Worchel and Mauney (4) study on the blind.

The wording of these instructions was given careful consideration for the vocabularies of most of our *Ss* were extremely limited. We endeavored to use the simplest terms possible. The *S* was allowed to read the directions thoroughly; then *E* went over each sentence with *S* and, where necessary, demonstrated in pantomime what was expected of him.

RESULTS

Tables 2 and 3 show the mean distance and *SD* of the first perceptions and final appraisals for every series of 30 trials and for every starting position, respectively, and the mean number of collisions with the wall, divided according to whether they occurred before or after the first perceptions.[2]

Collisions.—For our deaf *Ss* collisions with the obstacle were frequent, ranging from 83 (40% of the trials) to 209 (99.9%). Five *Ss* collided with the wall in more than 74% of the trials. Most of the collisions occurred prior to the first perception. Collisions occurring before the first perception indicate a lack of "obstacle-sense." Of course, if all these collisions occurred at the beginning of the learning series, and few, if any, toward the end of the series, then we might conclude that learning had occurred. This is not the case, however, since the mean number of collisions before "first perception" in the last series was almost double the mean number of collisions after the first perception. None of our deaf *Ss* attained the criterion of five or fewer collisions in any

[2] Data for the individual *Ss* in this experiment may be procured from the American Documentation Institute, 1719 N Street, N. W., Washington 6, D. C., by ordering Document 3561, remitting $1.00 for 35 mm. microfilm or $1.05 for 6 × 8 in. photocopies.

TABLE 3

MEAN DISTANCE (IN FEET) OF "FIRST PERCEPTIONS" AND "FINAL APPRAISALS" AND MEAN NUMBER OF COLLISIONS FOR EVERY STARTING POSITION FOR DEAF ($N = 10$) AND HEARING ($N = 5$) *Ss*

Ss	Report	Starting Distances from Obstacle (in Feet)											
		6		12		18		24		30		36	
		Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Deaf	First perception (p)	1.84	0.43	2.82	1.03	5.40	2.14	6.93	2.61	10.17	3.32	13.81	4.42
	Final appraisal	0.99	0.59	1.36	0.91	4.08	1.74	4.96	2.03	7.88	2.60	10.32	4.01
	Collisions												
	Before (p)	23.6	6.21	21.0	8.01	18.0	8.96	14.9	8.20	12.1	7.95	11.1	9.03
	After (p)	5.5	3.76	6.5	2.75	7.3	4.00	6.6	4.58	7.0	4.92	6.5	4.21
Hearing	First perception (p)	3.31	1.24	5.30	1.69	7.00	.96	11.60	2.73	12.41	3.32	15.04	2.21
	Final appraisal	.64	0.22	1.02	1.52	1.36	.55	2.08	1.10	2.95	1.89	3.12	1.74
	Collisions												
	Before (p)	0.2	0.97	0	—	0	—	0.2	0.97	0.2	0.97	0	—
	After (p)	2.2	1.33	2.6	1.01	1.2	1.02	1.4	0.80	0.6	0.80	1.4	0.73

one series. When the collisions are distributed according to the starting position (Table 3), the largest number is found to occur among the trials from the 6-ft. position and the smallest number from the 36-ft. position. Out of 35 trials at every starting position, from shortest to longest, the mean frequencies of collision for the deaf *Ss* were 23.6, 21.0, 18.0, 14.9, 12.1, and 11.1. It would be ridiculous to assume that our deaf *Ss* possessed the ability to perceive obstacles and that this ability functions best when the *Ss* start far away from the obstacle and poorest or not at all when they start close to it. A more tenable interpretation is that our deaf *Ss* were guessing, and by chance alone, they would collide more frequently with the wall at the shorter starting distances and less frequently at the longer, and by chance alone the frequency of collisions would vary inversely with distance.

The mean number of collisions for the deaf *Ss*, however, does show a progressive improvement during the experiment, for the number of collisions decreased throughout the learning series. Table 2 shows that from the first to the last series of trials, on the average, the deaf *Ss* collided with the obstacle 20.8, 16.2, 15.8, 14.3, 13.0, 9.9, and 10.6 times before the first perception. A reduction in the number of collisions can occur if *S* merely learns that he can avoid hitting the obstacle by stopping sooner. A more rigorous criterion than number of collisions is required as evidence of the perception of obstacles. As noted earlier, in almost every series, the collisions before "first perception" were double the mean number of collisions after "first perception," that is, *S* had no perception of the obstacle at all in these trials.

The performances of the hearing *Ss* who served under identical conditions as the deaf *Ss* present an entirely different picture (Table 2). No more than three series were required to attain a criterion of 25 successes in 30 trials. The maximum number of collisions was 14 for one *S*; the minimum was 5, for another *S*. None of the *Ss* gave any collisions before the

first perception in the second or last series. When the collisions are distributed according to the starting points (Table 3), we find no such relationship as was found with the deaf *Ss*. Collisions are distributed randomly throughout the various distances. For almost all the hearing *Ss* there is a sudden drop in the number of collisions after the first series. The mean number of collisions after the first perception dropped from 5.0 in the first series to 1.8 in the second and 1.6 in the third series. There were no collisions before "first perception" in the second and third series.

First perceptions.—The SD's of our deaf *Ss*' mean "first perceptions" show considerable variation from series to series. For our hearing *Ss*, there is much less variability from *S* to *S*. Individual differences are certainly expected, but it is rather surprising how little these differences are for the hearing *Ss*. Also, for both the hearing and deaf *Ss*, the mean distance for the "first perception" increases as the starting distances increase (Table 3). This may be explained by the fact that with more distance, *S* can perceive the obstacle sooner or he guesses at the "first perception." The latter hypothesis is more tenable for our deaf *Ss* because of the large variation noted from series to series and the large number of collisions before first perception.

Final appraisals.—The mean "final appraisals" for the deaf *Ss* in the seven series of trials are all large, varying from 5.14 ft. to 7.12 ft. This is the point where the deaf *S* approached the wall "as closely as possible without touching it." This certainly is not evidence for the ability to perceive obstacles. These results might mean, however, that the *Ss* possessed the ability to perceive the obstacle, but in a low degree. Even this conclusion does not seem warranted because the *SD*'s are extremely high, varying from 2.82 ft. to 3.72 ft. Also, if learning occurred, the averages of the final appraisals would decrease with practice, but this is not evident. Examining the results from series to series, we note some fluctuations, but, in general, no change in the ability to get "as close to the obstacle as possible without touching it." For example, from the first to seventh series successively, the mean final appraisals were 6.77, 7.12, 5.87, 5.14, 6.33, 6.19, and 6.91 ft. The fractionation of the data and the separate computation of the averages for every starting point reveal (Table 3), as with the number of collisions, that the distance of the "final appraisals" from the obstacle increased with the distance at which *S* started walking toward the obstacle. From the 6-ft. to the 36-ft. distance, the mean final appraisals were .99, 1.36, 4.08, 4.96, 7.88, and 10.32 ft. These facts indicate that the deaf *Ss* did not possess the ability to perceive the obstacle, and whatever perceptions were reported resulted from guessing. This is consistent with the findings of the earlier study with deaf-blind *Ss* (**3**).

The results of the hearing *Ss* are strikingly different from those of our deaf *Ss*. For every hearing *S* there is a consistent decrease in the average "final appraisals" from the first to the last series (Table 2). The mean for the hearing *Ss* on the third series (.90) is obviously reliably lower than the mean for the deaf *Ss* on their seventh and last series (6.91). A fractionation of the data and separate computation of the averages for every starting point (Table 3) reveal, however, that the final appraisals for our hearing *Ss* increased with the distance at which the *Ss* started walking toward the obstacle.

These results would apparently

lead us to reject the conclusion that the hearing *Ss* learned to perceive the obstacle. The fractionation of the data as presented in Table 3, however, includes all the data from the first through the last series. If in only one series, the "final appraisals" are perfectly correlated with the distance at which *S* started walking toward the obstacle, then, even though all the other series show zero correlation, the lumping together of all the "final appraisals" will result in a perfect relationship with starting points. Thus it is probable that during the early trials, the *Ss*, by chance alone, would stop farther away from the obstacle to report a final appraisal when the obstacle was set at the greater distances. That this is true is shown by the mean "final appraisals" of the first and last series. In the first series, the mean "final appraisals" for the hearing *Ss* are almost perfectly correlated with the starting points, that is, from the 6-ft. position to the 36-ft. position, the mean final appraisals are .75, 1.16, 2.14, 3.43, 4.31, and 6.96 ft. In the last series there is no such relationship: .41 ft. at 6 ft., .58 ft. at 12 ft., 1.00 ft. at 18 ft., .93 ft. at 24 ft., 1.07 ft. at 30 ft., and 1.15 ft. at 36 ft.

S's comments and behavior.—During the course of the experiment and after completion of each series, *Ss* were frequently questioned regarding the basis of their reports, and their overt behavior was closely observed. Most of the deaf *Ss* replied to questions by "do not know," "can't tell where board is," or "just a guess." All normal *Ss* showed considerable insight into the basis of their "final appraisals," but several could not explain how they made their "first perceptions." Several also reported a peculiar sensation of "blackness" as they approached the obstacle.

All normal *Ss* exhibited the same general type of behavior during the experiment and reported the use of auditory cues. They stamped and scraped their feet vigorously as they walked, in an effort to make as much noise as possible. Generally, they took several steps, stopped and stamped, then took several more, and stopped and stamped again. At the same time they listened for "changes in the quality of sounds." One *S* (LB) snapped his fingers in addition to stamping. Another *S* (JD) stamped his feet as soon as he was led up to the starting position, and frequently made "first perceptions" without taking a step. Occasionally, the noise made in leading these two *Ss* up to the starting position was sufficient for them to detect immediately the obstacle's presence. When they felt that the board was near, they stamped from one side of the path to the other. One of these *Ss* (LB) stated that he could hear "a change in the sound of the echo" as he stamped past the side of the obstacle. All normal *Ss* stopped walking when they felt that they were near and, instead, began to tap the right foot. At the same time they moved the foot slowly forward until it extended beneath the obstacle. This tapping generally enabled them to make "final appraisals" within inches of the board.

There was some disagreement over the type of change in the sound that occurred as the board was approached. According to Subject LB, "the echo comes back faster as I get closer to the board and it also gets louder. When I stamp to one side of the board and when the foot is extended beneath it, I no longer hear the echo." Subject JD said, "I hear more vibrations as I get nearer the board and the intensity of the sound increases. When my foot goes under the board, the sound disappears. That is, it sounded almost like when I started." Subject DT seemed to think that the sound became "higher" as he approached the obstacle, and that when he tapped his foot beneath the board the sound changed to a low thud. Subject MZ agreed that the sound changed, but he could not describe the nature of the change. At the beginning of the experiment Subject HH stated that "you feel the change (of sound) in your face. It is as though a pressure were being exerted on my face, and at the same time things get blacker and darker." Near the close of the experiment, however, she was more under the impression that she "heard the difference in sounds." The sensation of "blackness" and the "feeling of being closed in" are possibly associative factors based upon the change in the quality of sound which had come to signify an obstacle. These experiences were persistent in normal *Ss* and were not reported by deaf *Ss*. All normal *Ss* complained of the noise made by passing automobiles and children playing on the school grounds. Apparently the only auditory cues utilized were those produced by the *Ss* themselves.

Discussion

The second Cornell investigation with deaf-blind *Ss* dealt with the crucial problem of determining whether the aural mechanisms found to be the basis of obstacle perception in the first Cornell study were auditory or cutaneous (3). The conclusions of this second investigation were based upon a select group of extremely handicapped *Ss*, and it was felt by the present investigators that the laboratory conditions imposed upon these *Ss* might have made it difficult for any obstacle perception to be displayed. The present experiment dealt with a much younger heterogeneous group of individuals possessing varying degrees of deafness. No experimental limitation was placed upon the conditions under which the learning series was administered. With these extensions, the performances of our deaf *Ss* were very similar to those of the deaf-blind *Ss* used in the earlier investigation. The average distances of the first perceptions of our deaf *Ss* varied directly with the distance of the starting points—a result similarly obtained with the deaf-blind *Ss*. The SD of the means of our deaf *Ss*' first perceptions followed a similar trend—as with the deaf-blind *Ss*.

The average distance of *Ss*' "final appraisals" was measured in feet even after 210 learning trials. It may be that outdoor conditions present many more confusing cues than a laboratory and make it more difficult to learn to perceive the obstacle. With more practice, perhaps, the "final appraisals" might be smaller. The fact, however, that these "final appraisals" varied directly with the distance of the starting points, and that this correlation holds even in the last series for the two *Ss* on whom there were sufficient data to draw any conclusions, suggests that learning to perceive the obstacle was not evident. The variations of the averages of our *Ss*' "final appraisals" show no improvement with practice, a result obtained with the deaf-blind *Ss*. The collisions varied in number with the distance of the starting points, and in no case was the criterion of 25 successes in 30 trials obtained.

Almost all *Ss* reported at various times that they were guessing. Only the results of one deaf *S* were not as clear-cut and consistent as the other deaf *Ss*. This *S* (EC) gave some evidence of learning in the decrease in the average final appraisal and in her general behavior, but she reported "hearing something sometimes." An examination of her audiogram reveals a loss of 15 db at 10,000 cy. in the right ear. Her performance is certainly in accord with the conclusions of Cotzin and Dallenbach (1). Using pure tones, they found that the frequencies responsible for the perception of obstacles are those above 8000 cy. It is entirely possible that when she did hear she did perceive the obstacle. This interpretation may easily explain her erratic performance but consistent improvement. If the intensity of sound present in the area was sufficient for her to hear all the time, her performance might well have indicated her ability to perceive the obstacle.

One might still suggest that the conditions under which the experiment was performed were such that learning to perceive obstacles was impossible, and thus the investigation was prejudiced towards a particular hypothesis. Under identical conditions, however, a group of hearing *Ss* learned in comparatively very few trials to perceive the obstacle. The results of this group stand in opposition to those of the deaf *Ss*. Every hearing *S* reached criterion. Consistent improvement occurred in the *SD*'s of the "first perceptions" and "final appraisals." The averages of the "final appraisals" in the last series were approximately the same for all starting points. It is true that our hearing *Ss* gave slightly larger "final appraisals" than those

Ss who demonstrated in the Supa, Cotzin, and Dallenbach study (2) their ability to perceive obstacles. This is probably due to the presence of so many more confusing cues in the outdoor environment, such as wind, temperature changes, odors, distracting noises, and with further practice *S* might very well disregard these extraneous cues.

Summary and Conclusions

In order to determine whether the perception of obstacles under normal outdoor conditions requires auditory stimulation, ten deaf subjects and five with normal hearing were given 30 trials daily in the perception of a movable obstacle. These series were continued until *S* gave 25 reports in any one series without coming into contact with the obstacle. In no case was the number of series extended beyond seven (210 trials). The results of the study lead to the following conclusions:

1. Deaf *Ss* are incapable of learning to perceive the obstacle.
2. All normal hearing *Ss* learned to perceive the obstacle.
3. Audition is a necessary condition for the perception of obstacles by the blind.

(Manuscript received June 27, 1951)

References

1. Cotzin, M., & Dallenbach, K. M. "Facial vision": The role of pitch and loudness in the perception of obstacles by the blind. *Amer. J. Psychol.*, 1950, **63**, 485–515.
2. Supa, M., Cotzin, M., & Dallenbach, K. M. "Facial vision": The perception of obstacles by the blind. *Amer. J. Psychol.*, 1944, **57**, 133–183.
3. Worchel, P., & Dallenbach, K. M. "Facial vision," perception of obstacles by the deaf-blind. *Amer. J. Psychol.*, 1947, **60**, 502–553.
4. Worchel, P., & Mauney, J. The effect of practice on the perception of obstacles by the blind. *J. exp. Psychol.*, 1951, **41**, 170–176.

FREQUENCY OF USAGE AS A DETERMINANT OF RECOGNITION THRESHOLDS FOR WORDS [1]

RICHARD L. SOLOMON
Harvard University

AND

LEO POSTMAN
University of California

The tachistoscopic presentation of words has been used in attempts to detect the role of personality variables in perception. In general, the procedure has been to present printed words, drawn from various relevant "meaning" classes, for very short periods of exposure; the *S*'s task is to identify the words. Recognition thresholds are determined by increasing either the duration of exposure or the intensity of illumination. These thresholds have been considered to be an index of perceptual sensitivity or ease of recognition for the stimuli. The thresholds have also been interpreted as measures of the amount of "stimulus material" necessary for elicitation of the correct response.

The results of such experiments have lent some support to the view that personality variables may be significant determiners of perceptual sensitivity to visually presented verbal stimuli. The demonstrable effects of such factors are not large in absolute terms but often appear to be internally consistent and statistically reliable. Evidence has been presented for the selective effects of personal values and needs on perceptual sensitivity as measured in the tachistoscopic situation (**6, 7, 11, 12, 16**). It has also been suggested that words which may arouse anxiety because of their emotional connotations have significantly higher thresholds than neutral words (**8**). There is, however, some evidence that anxiety arousing stimuli may have lower thresholds than words to which *S* is indifferent (**2, 5**).

It is important to ask to what extent the apparent effects of complex personality factors reduce to the operation of basic psychological variables of wider generality. In a related series of experiments (**3, 12, 13, 14**) an attempt has been made to explore systematically some of the variables whose contribution must be taken into account in analyzing the effect of personality factors on recognition thresholds. Since the tachistoscopic studies of word recognition involve the presentation of verbal stimuli and the elicitation of verbal responses, these experiments concentrated on variables which have been shown to be extremely important in the psychology of *verbal learning*. Specifically, they have considered the effects of frequency and recency on the probability of occurrence of verbal responses in the presence of tachistoscopically exposed verbal stimuli. Howes and Solomon (**3**) have shown that when word length, word structure, and practice effects are roughly controlled or held constant, a very powerful correlate of the visual recognition thresholds for words is the *relative frequency* with which a word is used in the English language. Using the Thorndike-Lorge word counts (**15**) as an index of relative frequency of occurrence, these *E*s found that visual duration thresholds for tachistoscopically presented word stimuli seem to approximate a linear, inverse function of the logarithm of frequency of usage. The correlations between the thresholds and relative frequency of occurrence of words were found to range from .60 to .90, dependent upon the attendant experimental conditions.

The variable of word frequency had not been taken into account in previous experiments re-

[1] Parts of this research were supported by the Laboratory of Social Relations, Harvard University. Other parts were carried out at the University of California.

lating recognition thresholds to such factors as personal value (11). Solomon and Howes (14) showed that a very considerable amount of the variance attributed to personal value could be accounted for in terms of the frequency variable. Postman and Schneider (12) confirmed these findings. The question is still open whether the residual effects of value, interest, and so forth, can be reduced to individual *S*'s deviations from the frequency of word usage in the general population.

Howes and Solomon (4) applied a similar critical analysis to McGinnies' evidence for "perceptual defense" against taboo words (8). His taboo words had a low frequency of occurrence in the English language when compared to his neutral words, and higher thresholds for taboo words should, therefore, be expected on the basis of relative frequency of usage. McGinnies (9), in rebuttal, argued that the Thorndike-Lorge word counts are derived from literature, not verbal usage in everyday living, and so are not representative. This point is, however, quite inappropriate when it is realized that tachistoscopic presentation of words is very much like a reading situation; and, if laws of stimulus generalization are at all valid, frequency of occurrence of words *in literature* may be expected to predict visual duration thresholds for word stimuli at least as well as will some index of actual verbal response frequency in conversation. This problem remains to be explored.

In addition to the studies which have related frequency of usage to visual recognition thresholds, an experiment of Postman and Solomon (13) has shown that, when relative word frequencies are equated, the *recency* of word usage is significantly correlated with duration thresholds for verbal stimuli. Other things being equal, the more recently a word has been exposed to *S*, the lower will be the recognition threshold for that word stimulus. The variables of frequency and recency are not, of course, entirely independent. The more often a word occurs relative to all other words, the more likely it is to have occurred more recently than other words, either in literature read by the *S* or in words uttered by the *S*.

There is one fundamental weakness in the studies which have demonstrated the importance of frequency and recency in the determination of recognition thresholds for words. The index of frequency of word usage which these studies employed—the Thorndike-Lorge word counts—is a *population index* and allows us to estimate in only very rough ways the relative frequency of word usage for individual *S*s. This method, of course, represents a lack of adequate control over the word frequency variable. In all probability, the relationship between the relative frequency of word usage and recognition thresholds has been underestimated owing to this imprecision. Experimental control of relative word frequencies will make a more precise statement of the relationship possible. Once this frequency variable has been experimentally controlled and its effect on recognition thresholds measured, we shall have an important baseline for the evaluation of emotional, valuative, and need-related variables.

With this point in mind, and with a view to the inadequacy of previous indices of word frequency, two experiments were carried out. In both these experiments, relative frequency of word usage was "built in" experimentally and then related to the recognition thresholds for the words.

Procedure

Experiment 1.—Five undergraduates at Harvard University participated in this experiment. The *S*s were misled as to the purpose of the experiment. When *S* entered the experimental room, he was given the following instruction:

> This is an experiment concerning the effectiveness of repetition in learning to pronounce strange words correctly. It has a direct bearing on the problem of reading words in a foreign language, as compared to hearing the words spoken. In addition, we are interested in knowing whether the relative effectiveness of the two kinds of learning methods depends on general reading ability.
>
> We are going to give you a deck of cards. On each card is printed a strange word. We would like you to look at each card carefully and then pronounce the word in the way it would be pronounced if it were a word in the English language. Proceed steadily from card to card, turning over each one after you have finished with it. Go right through the deck and then stop unless you have serious doubts about your pronunciations.

TABLE 1

STIMULUS WORDS USED IN THE EXPERIMENTS

Core-Words		Filler-Words	
JANDARA	IKTITAF	KADIRGA	UDIBNON
AFWORBU	SARICIK	ADAFNAW	DILIKLI
BIWOJNI	ZABULON	BORULCE	MECBURI
NANSOMA	CIVADRA	NIJARON	OLMADIK
ENANWAL	LOKANTA	ENSHIMI	BOZULMA
		INKULAM	AMARIJA
		TAVHANE	FEVKANI

The *E* then handed *S* a pack of 100 cards. The pack consisted of a series of pronounceable nonsense words, seven letters each, typed in capitals across the middle of each card. The 100 cards contained 24 different nonsense words, repeated with varying frequencies. There were two words replicated 25 times, two words replicated 10 times, two words replicated five times, two words replicated 2 times, and 14 words appearing only once.

The pack of 100 cards was shuffled thoroughly before *S* went through it, card by card, reading and pronouncing each nonsense word. The nonsense words, selected from a Turkish-English dictionary, are presented in Table 1.

The first ten words of the list were the *core-words;* they were used later, in the tachistoscopic presentation. The other 14 words were "padding" to the deck, and they appeared one time each in the deck of cards. Each of the *core-words* was varied in its frequency of appearance in the pack from *S* to *S*, in such a way that a latin-square experimental design was completed.

This design was used in order to try to control for inherent differences among words in memorability and structural variations. Thus, each of the five *S*s was exposed to the same 24 words (ten core-words plus 14 padding words), but each *S* received different frequencies of exposure to the different core-words. Two core-words were used at each frequency, in order to replicate for words per se.

After the *S*s had read the 100 cards, they were given an irrelevant task in order to cover up the purpose of the experiment. *S*s were handed a book and given the following instructions:

> Now, we would like you to read the passage marked off in this book. We want you to read as clearly and distinctly as possible without faltering, in order that we may get a general index of your reading ability.

The *S* then read a page from a section on ethics in a randomly chosen textbook in philosophy. After this, *E* ushered *S* into another room for the tachistoscopic phase of the experiment and read these instructions:

> We are going to present to you, one at a time, some words. If you look in the eyepiece in this box, you will be able to see two lines. The words we will show you will appear directly between those lines. Each word will be presented to you for very short intervals of time, and at first you will not be able to tell what the words are. However, after each presentation we would like you to make a guess as to what the word was. Each word will be presented to you several times until you have correctly recognized it. I will say "Ready" before each exposure of the word, and I will tell you when you have correctly recognized the word and when a new word is going to be presented to you.

The words used in the tachistoscopic procedure contained the core-words of the pack of cards. These words were "buried," in random sequence, among 20 other words in the list of words for tachistoscopic presentation. These 20 words contained ten English words, varying in frequency over a large range in the Thorndike-Lorge counts, plus ten pronounceable nonsense words that previously had never been exposed to *S*. Thus, the total number of words to be exposed was 30. They were (a) the ten core-words; (b) English words (*welfare, surmise, testify, titular, example, promise, machete, venture, deserve, balance*); and (c) nonsense words never exposed before (*afsanaf, eksimek, peyapey, awnahdi, bilogos, akliyat, pisirik, nobetiki, levahik, fedakar*). In addition to these 30 words, four practice words were given at the start of the tachistoscopic session (*Harvard, teacher, university, professor*). The words in the tachistoscope were typed in the same capitals used on the cards of the deck described above. The ascending method of limits was used to determine the duration threshold (to the nearest .01 sec.) for each of the 30 words. The duration threshold was defined as the flash duration just necessary for correct recognition. For details of apparatus and procedure, the reader is referred to previous experiments (3, 11).

Experiment II.—The purpose and design of the second experiment were similar to those of the first experiment with the following exceptions: (a) The second experiment was run as a group experiment. (b) The *S*s—30 undergraduate and graduate students at the University of California—were instructed to pronounce the words subvocally after *E* had read them. (c) Because this was a group experiment, there was, of course, no possibility of counterbalancing the words, and, therefore, different frequencies were assigned to different words arbitrarily. This

procedure confounded frequencies and whatever effects stemmed from word structure. (d) Tachistoscopic exposure was by means of a projector equipped with a shutter which was set at a constant exposure duration of .01 sec. Each word was exposed 18 times with the illumination intensity increasing from exposure to exposure by an amount equivalent to a change of 2.5 v. (e) The number of trials required for the first correct recognition of the exposed word was used as a measure of the recognition threshold.

Results

Experiment I.—The average duration thresholds, in seconds, for all *core-words* are given in Table 2 and plotted in Fig. 1. The control words to which *Ss* had not been previously exposed, are listed as having a frequency of zero. Clearly, thresholds vary inversely with frequency of prior usage. Before evaluating the differences among the thresholds, we must take account of practice effects. The core-words were scattered through the tachistoscopic list randomly, but the same order of presentation was given to each *S*. In order to adjust for practice effects, the list was divided into thirds, and separate standard scores were computed for each third of the list. The standard scores for the three thirds were then comparable to each other, and further analysis could then be made on the assumption

TABLE 2

Average Recognition Thresholds for Core-Words of Different Frequencies

Core-Word Frequency	Experiment I		Experiment II	
	Av. Threshold (sec.)	Av. Standard Score	Av. Threshold (no. trials)	Av. Standard Score
0	.96	+.52	12.80	+1.02
1	.40	+.16	7.88	−0.24
2	.27	+.29	6.95	−0.48
5	.26	−.02	9.63	+0.09
10	.20	−.27	7.25	−0.44
25	.12	−.48	5.92	−0.77

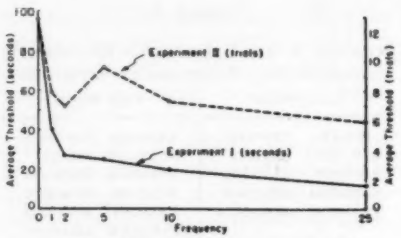

Fig. 1. Average recognition thresholds for core-words of different frequencies

that practice effects could not differentially affect thresholds for different core-words. The average standard scores for words with different frequencies of usage are listed in Table 2.

An analysis of variance was performed on the standardized threshold data (for core-words only). The main experimental variable—frequency—was found to be a significant source of variance ($F = 4.75$, $df = 4$ and 25, $p < .01$). Differences among sets of words fell short of significance ($F = 1.51$, $df = 4$ and 25). Finally, individual differences among *Ss* were significant ($F = 23.44$, $df = 4$ and 25, $p < .01$).

Experiment II.—The average duration thresholds, in number of trials required for recognition, are listed in Table 2 and presented graphically in Fig. 1. Table 2 also shows the average duration thresholds in standard score form. Analysis of variance again showed frequency to be a highly significant source of variance ($F = 45.60$, $df = 5$ and 145, $p < .01$). In the case of Exp. II, individual differences among *Ss* were not significant ($F = .32$).

Although the general trend of the thresholds as a function of frequency is the same as in Exp. I, the reversal between words having two and five presentations must be noted. It will be remembered that in Exp. II the core-words could not be counterbal-

anced among the various frequencies. It is probable that the particular structural characteristics of the words given two and five presentations outweigh the differential effect of frequency. When the Thorndike-Lorge word counts are used, the structural characteristics of particular words probably obscure the effects of frequency in a similar manner. In the absence of control for word structure, the general resemblance of the curves obtained in Exp. I and II is striking evidence for the importance of frequency of usage as a determiner of recognition thresholds.

Discussion

The data presented in Fig. 1 indicate quite clearly that visual recognition thresholds for words are a negatively accelerated decay function of frequency of prior exposure to a given word. In the previous work of Howes and Solomon (3), using the Thorndike-Lorge count as a word-frequency estimate, recognition thresholds have appeared to be an inverse linear function of the logarithm of frequency of word usage. A logarithmic transformation is not, however, appropriate to the present data. Whereas Howes and Solomon never had occasion to assign a frequency of zero to any English words, there is a true zero frequency point in the present function. The logarithm of zero frequency is relatively meaningless for our purposes without some sort of correction. The shape of the functions in Fig. 1, moreover, does not yield a linear plot when transformed logarithmically.

It is interesting to note, however, that the mean thresholds for frequencies of word usage varying from 1 to 25 (omitting the zero frequency) do yield a very good fit to an inverse linear logarithmic function. Therefore, there is some reason to suspect that, when a word has been exercised at least once, or has been encountered at least once, its recognition threshold will be roughly an inverse function of log frequency of prior exposure or usage.

The results shown in Fig. 1 provide a theoretical bridge between learning theory and perception theory. Let us assume that, by differentially exercising pronounceable nonsense words in this experiment, we have established associations of differential strength between the visually presented verbal stimulus and the response of reading or saying the word.[2] The strength of this association is measured by the duration or intensity of the stimulus exposure necessary for a correct verbal report.

Given a population of associations, the one which has been exercised most frequently will have the greatest probability of being elicited relative to other, like associations. How will this fact influence *S*'s responses in a tachistoscopic situation? When a stimulus pattern is presented at short durations or at low illumination intensities, only fragments of the total word stimulus are "effective." Such a stimulus fragment may be considered to represent a point on the generalization dimension of stimulus patterns capable of eliciting the correct verbal response. A given stimulus fragment may, of course, be located on several generalization dimensions, each involving a different word. Which verbal response will be given depends on the relative strengths of association which have been established, through generalization, between the particular stimulus fragment and the different response words. If the visually presented stimulus word has had a greater fre-

[2] For purposes of the present analysis it probably makes very little difference whether we think of these associations in terms of S-R connections or S-S patterns (hypotheses).

quency of prior usage than any of the competing response words, a correct response is highly probable.

Words of lower prior exercise frequency will be interfered with by words of higher exercise frequency. This interference will manifest itself in the tendency for *S*'s "guesses" to be high frequency words. If the actual stimulus word is a low frequency word, effective stimulus fragments will elicit erroneous "guesses" until the amount of effective stimulation becomes great enough on successive exposures to reduce the number of competing word responses. One may describe the increase in effective stimulation as limiting the range of competing "hypotheses" (1, 10) or one may speak of a restriction of stimulus generalization. In this connection it is interesting to point to the parallel between overt intrusions in retroaction and proaction experiments and wrong pre-recognition responses in the tachistoscopic situation. In both cases, a strong competing response temporarily replaces the correct response.

Frequency or prior occurrence used as an independent variable contains two aspects: (a) frequency of prior stimulation; and (b) frequency of prior response. When the Thorndike-Lorge word counts are used as a frequency index, these two aspects of environmental history are confused. Such is also the case in the experiments reported in this paper. When the pronounceable nonsense words were exposed to the *S*s, and they were asked to read them, either aloud or subvocally, a frequency index was being established into which the effects of exposure and response enter to unknown degrees. At present, we do not know whether frequency of prior *exposure* or frequency of prior *response* is the more important determinant of visual duration thresholds to words. A series of experiments aimed at this problem would have considerable bearing on the relative strengths and weaknesses of cognitive learning theory as compared with S-R learning theory in handling the facts of word-recognition experiments.

Our experiments point to frequency of prior usage as an important determinant of recognition thresholds but leave open the question of the process mediating the effect of frequency. The interaction of motivational factors with frequency has just begun to be explored (12, 14). Frequency of one type of sequence need not be equivalent to another type. The functions shown in Fig. 1 were obtained with verbal sequences which had no important motivational consequences for the *S*s. To what degree consequences of a high motivational nature will alter these functions is a question to be answered experimentally.

One final methodological note. Using operations typical of perception experiments (determination of thresholds), we have attacked our problem in terms of concepts derived from verbal learning. The phenomena of word recognition can play a strategic role in the rapprochement of theories of perception and verbal learning. The fact that many perceptual experiments concerned with nonverbal stimuli use verbal reports as their basic data makes it highly desirable to explore further the relations between the laws of verbal behavior and the laws of perception.

Summary

This experiment was concerned with the relationship between tachistoscopic recognition thresholds for words and frequency of prior usage of the words.

Pronounceable nonsense words were used as stimuli. Frequency of usage was controlled experimentally by requiring *S*s to read and pronounce different nonsense words with frequencies ranging from 1 to 25. Later, *S*s' tachistoscopic recognition thresholds for these words were determined as well as for control words which had zero frequency of prior usage. Recognition thresholds were found to vary inversely with frequency of prior usage.

The discussion stresses the close relationship between *Ss*' responses in a tachistoscopic recognition situation and responses measured in experiments on verbal learning. Frequency of past usage is a determiner of response strength in both types of situation. The importance of controlling such variables as frequency of usage in experiments relating perceptual sensitivity to personality variables is pointed out.

(Manuscript received June 30, 1951)

References

1. Bruner, J. S., Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald Press, 1951.
2. Bruner, J. S., & Postman, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, **16**, 69–77.
3. Howes, D. H., & Solomon, R. L. Visual duration threshold as a function of word probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
4. Howes, D. H., & Solomon, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229–234.
5. McCleary, R. A., & Lazarus, R. S. Autonomic discrimination without awareness: an interim report. *J. Pers.*, 1949, **18**, 171–179.
6. McClelland, D. C., & Atkinson, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, **25**, 205–222.
7. McClelland, D. C., & Liberman, A. M. The effect of need for achievement on recognition of need-related words. *J. Pers.*, 1949, **18**, 236–251.
8. McGinnies, E. M. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
9. McGinnies, E. M. Discussion of Howes' and Solomon's note on "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 235–240.
10. Postman, L. Toward a general theory of cognition. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads*. New York: Harper, 1951.
11. Postman, L., Bruner, J. S., & McGinnies, E. M. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, **43**, 142–154.
12. Postman, L., & Schneider, B. M. Personal values, visual recognition and recall. *Psychol. Rev.*, 1951, **58**, 271–284.
13. Postman, L., & Solomon, R. L. Perceptual sensitivity to completed and incompleted tasks. *J. Pers.*, 1950, **18**, 347–357.
14. Solomon, R. L., & Howes, D. H. Word frequency, personal values and visual duration thresholds. *Psychol. Rev.*, 1951, **58**, 256–270.
15. Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teachers' Col., Columbia Univer., 1944.
16. Vanderplas, J. M., & Blake, R. R. Selective sensitization in auditory perception. *J. Pers.*, 1949, **18**, 252–266.

ERRORS OF VISUAL RECOGNITION [1]

F. H. GEORGE
University of Bristol

Gestalt psychologists maintain that in memory function certain dynamic changes take place in the "memory trace"; these processes are labeled variously normalizing, sharpening, pointing, etc. These views, now well known, have had a considerable effect on the development of experimental psychology, and they are based largely on the experimental work of Perkins (**10**), Wulf (**12**), and Gibson (**2**). However, Hebb and Foord (**3**) have recently challenged the validity of the method of successive reproduction which was used by all the above three *E*s. Hebb (**4**) has recently made this challenge a nodal point in the development of his own theory of perception, hence the renewed importance of investigating further the empirical facts.

There is one crucial question that needs to be asked. Is the method of successive reproduction valid to show these dynamic changes in the memory trace? Hebb says "No" and we shall agree that the method is, in fact, open to serious objections. It is clear that in the course of successive reproductions simplification may take place owing to *S*'s "remembering" previous reproductions. It may be counter-objected that simplification of a certain symmetrical or normalizing kind even under these conditions shows certain dynamic changes. This appears to be a matter open to further experimental check and such a check has been carried out by Hebb and Foord with the method of a single reproduction. The argument here is simply that the questionable method of successive reproductions is avoided and any progressive dynamic changes should be observable in the one reproduction. Under these new conditions Hebb and Foord found no statistically significant progressive change in the trace. Thus it would appear to leave the Gestalt psychologists without any valid evidence in favor of their principle. However, further difficulties have arisen, as unfortunately there are points in this experiment which are not beyond criticism. Hebb and Foord used 5-min. and 24-hr. time intervals between the original stimulus situation and the time of reproduction. Is it possible that the changes may have been progressive if the time interval had been longer? Secondly, Hilgard (**5**) has pointed out that the figures originally chosen may have been equally balanced between "sharpness" and "closure." Two experiments were carried out in an attempt to overcome the two possible objections to the Hebb-Foord experiment.

[1] I would like to acknowledge the considerable help I have received from Mr. B. Waldron and Miss A. Passmore of the University of Bristol, in carrying out this experiment.

METHOD

In Exp. I, the sets of cards originally used by Hebb and Foord (**3**) were used, under identical (or near identical) conditions of instructions, etc., but with new time intervals. A control group was used having the same 5-min. and 24-hr. intervals as in the original work, but new groups were also set up having time intervals of 4 and 8 days.

Experiment II was under the same experimental conditions as the Hebb-Foord experiment, with time intervals of 5 min. and 24 hr., but with a set of ellipses of varying eccentricity.

Stimulus materials.—Three sets of cards were used altogether. The first set was composed of 24 circles which were incomplete, there being small variable gaps in the circumference.

These gaps varied by steps of 2½° up to a 60° gap which was the maximum. The set of incomplete circles was drawn with India ink on separate 7.5-cm. squares of white bristol board. Each circle had a diameter of 5.6 cm. Cards 8 and 16 were duplicated and used as original stimulus cards; of these duplicated cards, one was selected and shown to *S* as the original stimulus. The *S* was given precise instructions before the presentation of the stimulus cards, and these will be remarked on later.

The second set was composed of 24 arrowheads in which the sharpness of the point of the arrow varied progressively. The distance between the barbs of the arrowheads was constant. The length of the horizontal line joining the point of the arrowhead to the centre of its notch was also constant: the set actually varied from a very pointed arrowhead with a deep notch to a laterally unsymmetrical quadrilateral. Cards 6 and 14 of this set were duplicated and used for presentation to *S* as the original stimulus. Both of the above sets of cards were the same as those sets used by Hebb and Foord (3).

The third set was composed of 14 ellipses of progressively varying eccentricity; the variation in shape was from nearly circular to flat and long. The eccentricity of an ellipse is defined by: $e = \sqrt{\frac{a^2 - b^2}{a}}$, where a and b are the semimajor and semiminor axes respectively. The values of e^2 varied from approximately .1 to .8. The actual values of e^2 may be calculated for the 14 ellipses since $2a$ varied from 3⅞ in. in Card 1 by steps of 1/12 in. to 4⅞ in. in Card 14, and $2b$ varied from 3⅛ in. to 2⅛ in. by steps of 1/12 in.

The cards were drawn in India ink on bristol board, as in the case of the first two sets, but in this case the cards were of larger dimensions, being 15 cm. by 14 cm. Two of the set, Cards 4 and 8 of eccentricity .61 and .76, were duplicated and used for presentation to *S* as the original stimulus. The line of the ellipses was approximately $\frac{1}{32}$ in. in thickness and this was included in the values of a and b.

Subjects.—The *S*s, 200 in number, were largely undergraduates of the University of Bristol, but some number of nonacademic civilians were employed.

Experimental design.—In Exp. I four groups of *S*s were used. The time intervals between original presentation of the stimulus cards and the re-presentation in these groups were 5 min., 24 hr., 4 days, and 8 days. The groups were then divided into halves and one half was shown the original "circle-with-gap" (Card 8) and either Card 6 or Card 14 of the arrowhead series. The other half was shown the other original "circle-with-gap" (Card 16) with either Card 6 or Card 14 of the arrowhead series.

In Exp. II, two groups were set up, one group having a 5-min. interval between original presentation and re-presentation, while the other had a 24-hr. interval. These two groups were also divided roughly into half and one group was shown the original ellipse Card 4 and the other half was shown Card 8.

Procedure.—Each *S* was treated individually. The *S* was shown a circle-with-a-gap card (Card 8 or Card 16) and then an arrowhead diagram (Card 6 or Card 14). In Exp. II, *S* was presented with one or other of the ellipses (Card 4 or Card 8). The cards were viewed for 5 sec., and in Exp. I the time interval between the presentation of the two stimulus cards was 20 sec.

The *S* was told that he would be shown a diagram which he was to remember exactly, and that the accuracy of his memory would be tested. At the end of the prescribed retention interval *S* was handed the 24 recognition patterns in Exp. I or 14 in Exp. II, each on a separate piece of bristol board as in the case of the original stimulus card, and told to pick out the one he had been shown originally.

The packs of cards were handed to *S* in a bundle and *S* was instructed to sort through them in order as many times as he wished, but keeping the cards in the original order and looking at only one at a time. As he discarded one card he placed it face downwards on the table to make a new pile and thus retain the cyclical order of the cards. The pack was handed to all *S*s in the same cyclical order but the top card was varied for each *S* in a random manner. Thus the series was interrupted at any possible point, and thus it was hoped to avoid a constant error by having, say, Card 1 always on top.

Results

The results seem quite conclusive. There is no progressive change observable in the trace, such as would be predicted by a theory of Prägnanz. This confirms the findings of Hebb and Foord over a wider range. It is seen that even if the time interval is extended up to 4 days and 8 days no progressive tendencies are observable. These results demand little comment as the statistical analysis is surprisingly conclusive. The second experiment is also consistent with the general findings. Using a set of ellipses

TABLE 1

ERROR DISTRIBUTIONS AND MEANS IN RECOGNITION OF FORMS AFTER VARYING INTERVALS OF TIME: EXPERIMENT I

Amt. of Error*	Circle No. 8				Circle No. 16				Arrow No. 6				Arrow No. 14			
	5-Min.	24-Hr.	4-Day	8-Day	5-Min.	24-Hr.	4-Day	8-Day	5-Min.	24-Hr.	4-Day	8-Day	5-Min.	24-Hr.	4-Day	8-Day
−13 to −11	—	—	—	—	1	—	—	—	—	—	—	—	—	—	—	—
−10 to −8	1	—	—	1	1	—	—	—	—	—	—	—	—	—	—	—
−7 to −5	1	1	1	2	2	—	3	3	—	1	—	—	1	—	1	1
−4 to −2	4	7	2	3	1	3	2	2	2	5	6	5	3	2	2	3
−1 to +1	7	2	3	5	4	3	4	2	10	7	4	5	7	9	8	5
+2 to +4	2	2	2	2	2	4	5	4	4	2	4	2	4	4	3	2
+5 to +7	3	1	4	1	2	5	2	4	—	1	3	3	2	2	1	2
+8 to +10	1	1	2	—	1	—	1	2	—	—	—	—	—	—	—	1
+11 to +13	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
+14 to +16	—	1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
N	20	16	14	14	14	15	17	17	16	16	17	15	17	17	15	14
Mean	0.90	1.31	2.57	−1.30	0.64	2.20	0.71	1.76	0.38	−0.56	0.70	0.60	0.50	1.06	0.20	0.86
Direction of change**		P	P	NP		NP	NP	NP		NP	P	P		P	NP	P

* A minus sign signifies an error in the direction of Card 1, where Card 1 has the minimum gap, in a set of *circles* numbered from 1 to 24. In the case of arrows the same system is employed, with Card 1 representing the most pointed arrowhead.

** P is for "progressive," and NP is for "non-progressive." The change is regarded as progressive if the mean error is greater and in the same direction after the stated interval than after 5 min.

under the Hebb-Foord conditions, no progressive change, indeed no significant change at all, is observable in the results.

Table 1 presents the distributions and means of errors of recognition in Exp. I, in terms of minus or plus deviations from the original stimulus. None of the differences between the 5-min. and longer intervals is significant at even the 5% level of confidence, as tested by *t*. Furthermore, if the change is regarded as progressive when in the experiment the mean error was greater in the same direction after the longer interval than after 5 min., in no case is there any sign of a significant progressive change.

Table 2 presents a similar analysis of the data on recognition of ellipticality that were obtained in Exp. II. Again there is no evidence for a significant progressive change in the error with elapsed time between the original presentation and the recognition test.

TABLE 2

ERROR DISTRIBUTION AND MEANS IN RECOGNITION OF ELLIPSES: EXPERIMENT II

Amount of Error*	Ellipse No. 4		Ellipse No. 8	
	5-Min.	24-Hr.	5-Min.	24-Hr.
−2	2	3	1	1
−1	3	5	4	2
0	4	3	8	5
1	8	2	2	4
2	3	2	4	4
3	2	—	1	—
N	22	15	20	16
Mean	0.59	−0.33	0.35	0.50
Direction of change*		NP		P

* See footnotes to Table 1.

DISCUSSION

The above experiments corroborate the Hebb-Foord (3) results over a wider range. They seem also to meet the original objection leveled against the Hebb-Foord experiment by Hilgard (5). Thus it may be repeated that the results fail to corroborate the Gestalt hypothesis of dynamic activity within the trace.

It should be added that it is not strictly necessary to demonstrate the lack of experimental evidence for an hypothesis for which no valid verification is available, although here it may be objected that in carrying out the above experiments the possibility of such traces is provisionally accepted. Clearly such experiments as the above do, in fact, serve to eliminate superogatory hypotheses in psychology. It further points to the need for introducing hypotheses that can be subjected to adequate experimental verification and thus to operational definition (1). It should be noted further that all possibility of any general principle of Prägnanz in memory has not vanished as a result of these experiments. It does however eliminate the possibility of any general principle of Prägnanz as so far postulated. Any future attempt to use such a principle must be accompanied by an operational definition in which the limits of application are stated clearly, thus allowing verification.

It should be added that in the search for conditions wherein the dynamic changes are possible, or take place, we can now tentatively eliminate two variables and narrow the limits of the search. The lengthening of time interval not being a relevant factor here, it would seem that there are some situations wherein dynamic changes do not take place and the length of time involved makes no difference. To make a bolder and experimentally unjustified inference, it seems that if dynamic traces occur they may occur immediately. Secondly, it would seem that the use of figures which are clearly not equally balanced between "closure" and "sharpness" is not a guarantee that dynamic changes take place. There is one further point of special interest to the theoretical psychologist. In the theory of Gestalt psychology the principles were largely constructed by analogy with the classical physics of closed energy systems: the principle of Prägnanz being a particular example of this sort of theoretical prediction. However, the present weight of evidence suggests that open energy systems are the only possible basis for biological fields and for psychological theory in particular, and many of the Gestalt analogies cannot therefore logically be "taken over" or derived in the way they have been. This means that much of Gestalt theory needs to be reformulated in terms of open energy systems: the work of von Bertalanffy (11) and Krech (7, 8, 9) as well as Hebb (4) therefore predicts the failure of the principle of Prägnanz on theoretical grounds. Thus, it seems that it must be dropped completely until it can be shown to take place and the conditions under which it can be shown to take place carefully defined.

There is one point in the experimental method that deserves mention. The *S* was told that he would be shown a diagram which he was to remember *exactly*, and that the accuracy of his memory would be tested. This is the Hebb-Foord wording and was retained word for word as a necessary part of the controls of the experiment. It is possible that the word "exactly" could be construed as effecting a specific suggestion on the *S*s and influencing the form of their recognition. However, while this remains possible, as a matter of clinical observation, the word in question was not emphasized and it appeared to have no special effect. Statements made subsequently by *S*s showed that they were unaware of any special influence of the above. They in fact tended to remember verbal phrases such as "incomplete circle," "black ellipse," etc., although it must be admitted that the question of suggestion involved in the instructions remains a consideration, however small.

Summary

Two experiments were undertaken on visual recognition, in which, in all, three sets of stimulus card were used, incomplete circles, arrows, and ellipses. Time intervals of 5 min., 24 hr., 4 days, and 8 days were used in the cases of circles and arrows, and 5 min. and 24 hr. for the ellipses. The findings were in accord with the Hebb-Foord findings, that there is no observable progressive change in recognition as it is predicted by the principle of Prägnanz. Thus, the principle of Prägnanz must be dropped altogether or redefined in more specific terms, the limits of application being clearly stated.

(Manuscript received July 11, 1951)

References

1. Bridgman, P. W. *The logic of modern physics.* New York: Macmillan, 1938.
2. Gibson, J. J. The reproduction of visually perceived forms. *J. exp. Psychol,* 1929, **12,** 1–39.
3. Hebb, D. O., & Foord, E. N. Errors of visual recognition and the nature of the trace. *J. exp. Psychol.,* 1945, **35,** 335–348.
4. Hebb, D. O. *The organization of behaviour.* London: Hamish Hamilton, 1949.

5. HILGARD, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
6. KOFFKA, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.
7. KRECH, D. Notes towards a psychological theory. *J. Pers.*, 1949, **18**, 66–87.
8. KRECH, D. Dynamical systems, psychological fields and hypothetical constructs. *Psychol. Rev.*, 1950, **57**, 283–290.
9. KRECH, D. Dynamic systems as open neurological systems. *Psychol. Rev.*, 1950, **57**, 345–361.
10. PERKINS, F. T. Symmetry in visual recall. *Amer. J. Psychol.*, 1932, **44**, 473–490.
11. VON BERTALANFFY, L. An outline of general system theory. *Brit. J. Phil. Sci.*, 1950, **2**, 134–165.
12. WULF, F. Veränderung von vorstellungen. *Psychol. Forsch.*, 1922, **2**, 333–389.

RETENTION OF VERBAL MATERIAL AS A FUNCTION OF MOTIVATING INSTRUCTIONS AND EXPERIMENTALLY-INDUCED FAILURE [1]

WALLACE A. RUSSELL

State University of Iowa [2]

The importance for learning of such factors as motivation and frustration has been widely recognized, yet the systematic investigation of the effect of these variables on human learning and retention has scarcely begun. In both experimental and clinical situations we frequently seek to control motivation and frustration by verbal communication between the psychologist and *S*. If the effects of such communications on specific aspects of performance related to learning could be more reliably determined, the possibility of constructing adequate systematic accounts of the behavior changes involved would be markedly improved.

This investigation deals with three questions relevant to these considerations.

First, what are the effects of verbally-induced failure upon retention in an experimental verbal learning situation? Preliminary studies by Farber, Russell, and Andreas (**2**) have demonstrated that in paired-associate learning failure adversely affects performance on immediate tests of retention, but that on delayed tests the failed groups perform as well as, or perhaps even better than, non-failed control groups. This interesting picture of initially depressed and subsequently improved performance led to the hypothesis that the experience of failure resulted in both an inhibitory effect and an increase in drive. The dissipation in time of the depressing inhibitory effect, it was assumed, allowed drive to assert itself in the relative improvement of retention scores in the failed groups on delayed tests. Since the delayed motivational effect had not been clearly established statistically, an effort was made in this study to provide more precise tests for this possibility. At the same time, the generality of the immediate disruptive effect of failure was tested by employing a verbal learning task differing from that used in the preliminary studies.

Second, what are the effects of initial motivating instructions upon the learning and retention of a verbal task? Efforts to control human motivation usually involve verbal instructions of some kind. Although such so-called motivating instructions do not seem to result in behavior changes completely analogous to those obtained by such procedures as depriving an organism of food, they have apparently been used effectively by Alper (**1**) and others. Her "task-oriented," or Low Motivation *S*s were merely told to memorize a list of nonsense syllables and numbers, while her "ego-oriented," or High Motivation *S*s were told that the task was a test of intellectual ability. An immediate test of recall showed no significant differences between the groups. A second test of recall on the following

[1] This article is based on a dissertation submitted to the faculty of the Department of Psychology of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Dr. I. E. Farber for his criticisms and suggestions during the planning and execution of this study.

[2] Now at the University of Minnesota.

day, however, showed a striking difference. The "ego-oriented" *Ss* showed no significant retention loss, whereas the "task-oriented" *Ss* dropped significantly below their first recall score. If motivational differences, as defined by verbal instructions, do not affect performance in original learning or immediate recall, but do result in differences in recall on delayed tests, important modifications of our theoretical treatments of forgetting may be demanded. This possibility was explored in the context of the present experimental situation by testing the influence of differential motivating instructions during learning and at recall intervals of 1 min. and 24 hr.

The third question involves the possibility of a relationship between the initial motivating instructions and the subsequent consequences of failure. No one now denies that the effect of failure may vary considerably according to the specific situation in which it is induced. The necessity for a more detailed investigation of the antecedent factors determining failure effects has seldom been made explicit, however, and experimental evidence on the problem is almost completely lacking. Rosenzweig (5), in a study using the Zeigarnik technique, found that recall of incompleted tasks seemed to vary with the type of motivating instructions. If incompleted tasks are considered as failures, this may indicate that the magnitude of failure effects is influenced by the level of motivation. This hypothesis was tested by determining whether there was any interaction between the level of motivation as defined by instructions and the magnitude of the subsequent effects of failure.

To summarize, then, this investigation undertook to determine the joint influences of motivating instructions and experimentally-induced failure upon immediate and delayed retention in a verbal learning task.

Method

Subjects.—The 128 *Ss* were students in introductory psychology during the spring (94) and summer (34) of 1949. All *Ss* were unpracticed in experimental verbal learning tasks. The 86 males and 42 females were assigned at random to the experimental groups.

Apparatus and Materials.—The materials to be learned were two lists of 12 three-letter nonsense syllables previously constructed by Melton (4). These were presented in serial order on a standard Hull-type memory drum, and learned by the anticipation method. Each syllable was exposed for 2 sec. A black shutter then covered the exposure window for two more seconds before the succeeding syllable was revealed. Thus, between the appearance of one syllable and another, there were 4 sec. during which *S* could endeavor to spell out the syllable that would be next exposed. The score for a given trial was the number of syllables correctly anticipated on that trial. A 6-sec. interval elapsed between each presentation of the list.

Procedure.—All *Ss* followed the same general sequence of treatments: After being given the conventional instructions for learning nonsense syllables on a memory drum, all *Ss* learned the first 12-syllable list (the practice list) to a criterion of six correct anticipations, with the instruction that the task was presented for the purpose of practice only. Then one of two sets of instructions intended to produce differential motivation was presented verbally to *Ss*, and then they learned a second 12-syllable list (the experimental list) to a criterion of five correct anticipations, and were then allowed one additional learning trial. Immediately following this learning each *S* was either told that he had failed at the experimental task or was given no knowledge about his performance. Finally, each *S* was given a retention test on the experimental list by the relearning method either 1 min. or 24 hr. after reaching the learning criterion. In either case, relearning was continued to a criterion of two consecutive trials without any errors in anticipation.

The 128 *Ss* were randomly assigned to eight groups of 16 *Ss* each. The eight groups represented all of the possible combinations of high and low motivating instructions, failure instructions or rest, and immediate or delayed recall. The treatment alternatives were as follows.

Motivating instructions.—One of the two sets

of motivating instructions was introduced just prior to the learning of the experimental list. Four groups received what were termed "H," or High Motivation, instructions. These indicated that performance on the task was a test of scholastic aptitude; the importance of each individual's performance was stressed, and an incentive of double credit[3] for serving was offered for a good score. The *S*'s name was ostentatiously written on his response sheet and he was told that his next task was "for the record."

The remaining four groups were given the "L," or Low Motivation, instructions. These presented the experimental task as another practice list. These lists, it was said, were being tried out for possible use in future research work, and the present purpose was stated to be the determination of which ones might be suitable for such use. This preliminary work was described as montonous but necessary, even though the scores obtained were not meaningful in themselves.

Failure-rest instructions.—After reaching the learning criterion on the experimental list, two groups from each of the H and L motivation conditions were given instructions intended to produce feelings of failure, while two other groups in each condition were engaged in neutral conversation. The "F," or failed groups, were told: "Mr. ——, you seem to be having trouble with your test. I don't understand what the matter is. You have been doing much worse than the other people who have worked on this test. So far you're failing badly. You have had——trials and yet you failed to get more than——right on the last trial. (*E* shows *S* his score sheet.) Most subjects get that many right in half as many trials. You're not even close to an average score for——trials. You're so far below the norm that I thought something must be wrong out here. (Pause) You can see all right, can't you? I don't understand why you've been so unsuccessful. So far this is one of the lowest scores we've had."

The "R," or Rest, groups were given no knowledge of their results but were engaged in irrelevant conversation for a period of time (1 min.) equal to that required for the induction of failure. It was announced that it was time for a rest, and *S* was asked what section of introductory psychology he was in, whether he had ever attended summer school, etc.

Measurement of retention.—The eight treatment groups were finally differentiated in terms of the time interval between learning and a test of retention by the relearning method. Four groups were given an immediate test only 1 min. after the end of original learning, while four comparable groups were dismissed following the Failure-Rest instructions and recalled 24 hr. later for a similar but delayed test of retention. Groups receiving the immediate test are termed "I" groups, and those receiving the delayed test, "D" groups. Following the relearning test all *S*s were told that they had performed quite satisfactorily and were excused from the laboratory.

The treatment accorded each of the eight groups is indicated by an identifying label, which consists of three letters. The first letter indicates the motivational condition; the second, the failure-rest condition; and the third indicates whether relearning was immediate or delayed. Thus, for example, group HFD received High Motivation instructions (H), was given instructions (F) following learning which stated that they had failed, and was tested (D) for retention and relearning after a delay of 24 hr.

For each *S* records were kept which indicated the correct anticipations, the incorrect anticipations, and the failures to anticipate for each word on all trials.

[3] The *S*s were taking a course which required its members to serve for 3 hr. as experimental *S*s. Double credit implied that 1 hr. of experimental participation would be counted as two.

Results

Performance on the practice list provided a basis for comparing the groups with respect to initial learning ability. A simple analysis of variance was carried out to test whether the mean number of trials required to reach a criterion of five correct anticipations on the practice list differed significantly from group to group. The value of F obtained in this analysis was less than unity, indicating that the means did not differ, and that the assumption of original equality of the groups, so far as rate of improvement was concerned, was tenable.

Initial Learning of the Experimental List

The experimental list was learned to a criterion of five correct anticipations. Table 1 presents the means and *SD*'s for the number of trials required to reach the criterion in all eight groups. Inspection of Table 1

TABLE 1

MEANS AND *SD*'S OF THE NUMBER OF TRIALS REQUIRED TO REACH THE LEARNING CRITERION

(*N* in each group = 16)

Group	Mean	*SD*	Group	Mean	*SD*
LRI	7.31	2.49	HRI	6.06	2.54
LFI	6.69	1.89	HFI	8.31	4.38
LRD	6.81	2.38	HRD	6.75	3.38
LFD	7.38	1.65	HFD	8.06	3.56

reveals that the *SD*'s for the four High Motivation groups (HRI, HFI, HRD, HFD) were all higher than any of the *SD*'s in the Low Motivation groups (LRI, LFI, LRD, LFD). Moreover, the variance for the pooled H *Ss* was 13.30 as contrasted with 4.64 for the L *Ss*. The *F*-ratio between these pooled variances was 2.87, a value significant beyond the 1% level of confidence for the *df* involved (63 and 63).

In spite of this difference in the variability of the H and L groups, there appeared to be no reliable difference between them with respect to the mean number of trials required to reach the criterion. The Mann-Whitney *U*-test (**3**), which is applicable in the case of unequal variances, showed that rank-order differences as great as those found among the four L means and the four H means would be expected by chance alone 44% of the time.

It was concluded that the two sets of motivating instructions did not differentially affect the mean number of trials required to reach the criterion on the original learning of the experimental list, but that in terms of this measure of performance at least, the High Motivation instructions did tend to increase the variability of performance more than the Low Motivation instructions.

Retention of the Experimental List

In order to compare the amount of retention in the various groups, the difference between the scores on the first relearning trial and the last trial of original learning (the trial following the achievement of five correct) was computed for all *Ss*. In terms of this difference, a positive score indicated an increment in performance on the first relearning trial relative to the last learning trial, and a negative score reflected poorer performance. This difference score reflects the direction and magnitude of change in performance scores rather than the absolute levels of performance, and will generally be referred to hereafter as the adjusted recall score. The means and *SD*'s of the scores for the last trial of original learning, the first trial of relearning, and the adjusted recall scores are presented in Table 2.

Immediate retention.—The data for the four I groups and the four D groups were analyzed separately. The adjusted recall scores for the groups receiving the immediate test were subjected to a double-classifica-

TABLE 2

MEANS AND *SD*'S OF CORRECT ANTICIPATIONS ON LAST LEARNING TRIAL AND FIRST RELEARNING TRIAL, AND ADJUSTED RECALL SCORES

(*N* in each group = 16)

Group	Score on Last Learning Trial		Score on First Relearning Trial		Adjusted Recall Score	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
LRI	5.69	1.40	6.94	1.64	1.25	1.64
HRI	5.44	0.99	6.38	1.22	0.94	1.30
LFI	5.75	1.64	5.19	2.06	−0.56	1.66
HFI	5.69	1.21	4.38	2.31	−1.31	1.10
LRD	5.12	0.86	1.50	1.27	−3.62	1.62
HRD	5.50	1.90	3.06	2.25	−2.44	2.42
LFD	5.81	1.07	2.56	2.06	−3.25	2.25
HFD	5.25	1.25	2.19	1.94	−3.06	1.78

tion analysis of variance, the summary of which is presented in Table 3.

It may be seen that the effect of the Failure-Rest variables upon immediate retention was highly significant. With 1 and 60 *df*, an *F* of 7.08 is needed for significance at the 1% level, whereas the obtained *F* for Failure-Rest effects was 29.71. The mean adjusted recall scores, shown in Table 2, reveal a drop in performance for both Failure groups (LFI and HFI) between the last trial of original learning and the first trial of relearning, in spite of the fact that only 60 sec. intervened between these two trials. Groups HRI and LRI, for which neutral conversation was substituted for the statements implying failure, showed an increment rather than a decrement in performance following the 60-sec. interval. The *t*-test[4] in scores of groups HFI and HRI, and between groups LFI and LRI, yielded values of 4.27 and 3.44, respectively. Both *t*'s are significant beyond the 1% level.

Relative to the performance of the Rest groups, then, the disruptive effects of failure upon immediate retention appear to have been clearly demonstrated.

This immediate disruptive effect of failure, however, appeared to be quite temporary. Performance on subsequent relearning trials indicated that the differences between comparable F and R groups diminished rapidly after the first trial. There was, however, no evidence to support the hypothesis that the F groups would ultimately perform better than R groups. The absence of any pronounced facilitating effect following the induction of failure was emphasized by appropriate *F*-tests which revealed that the groups did not differ significantly in the number of trials required to reach the relearning criterion of two consecutive trials without errors.

[4] In this *t*-test, and others cited hereafter which deal with differences between groups appearing in the same double-classification analysis of variance, the individual's variance served as the estimate of the population variance.

TABLE 3

ANALYSIS OF VARIANCE OF ADJUSTED RECALL SCORES IN GROUPS RECEIVING THE IMMEDIATE TEST OF RETENTION

Source of Variation	*df*	Sum of Squares	*F*
Failure-Rest	1	66.01563	29.71*
Motivation (H vs. L)	1	4.51563	2.03
F × M	1	.76562	.34
Individuals	60	133.31250	
Total	63	204.60938	

* Significant beyond the 1% level of confidence.

The analysis of variance of the adjusted recall scores for the I groups does not reveal any significant effect attributable to the High and Low Motivation instructions (see Table 3). The *F*-test for High vs. Low Motivation effects yielded a ratio of 2.03. This falls far short of significance at the 5% level. In the case of the original learning, the High Motivation groups did appear to be more variable, in terms of the number of trials required to reach the criterion. It may be seen from Table 2 that the effect of the two motivation conditions upon the variability of the various scores computed in terms of the number of correct responses was both slight and inconsistent.

The interaction between the failure and motivation conditions was not significant. At least in this study, there is no indication that the magnitude of failure effects varies with the motivational instructions given prior to learning. This, of course, could be due to the particular instructions

TABLE 4

ANALYSIS OF VARIANCE OF ADJUSTED RECALL SCORES IN GROUPS RECEIVING THE DELAYED TEST OF RETENTION

Source of Variation	*df*	Sum of Squares	*F*
Failure-Rest	1	.2500	.06
Motivation (H – L)	1	7.5625	1.70
F × M	1	4.0000	.90
Individuals	60	267.6250	
Totals	63	279.4375	

used, or to the aspect of behavior (retention) measured.

Delayed retention.—An analysis of variance, comparable to that carried out for the immediate retention groups, was performed on the adjusted recall scores of the four delayed retention groups. The summary of this analysis is presented in Table 4.

An *F*-test for the significance of the effect of the Failure-Rest variable upon delayed retention resulted in a value of .06. Since this value is less than unity, it was concluded that the Failure-Rest conditions did not differentially affect delayed retention.

An examination of Table 2 reveals that all four delayed retention groups showed decrements in recall on the first relearning trial. Since the smallest decrement, that of 2.44 for group HRD proved reliable beyond the 1% level with a *t* of 3.27, it follows that all the other decrements were also significant beyond that level.[5] Moreover, since the greatest difference between any of the four means, that of 1.18 between groups HRD and LRD, yielded a *t* of 1.59, significant at the 10% to 20% level, it also follows that differences between any of the other means were even less reliable.[6] It can therefore be stated that the decrements of 3.25 in group LFD and of 3.06 in group HFD both represented significant drops in performance on the first relearning trial, but that these effects do not differ significantly from each other or from the corresponding decrements in groups HRD and LRD. These latter statements, of course, elaborate the conclusion that delayed retention was not affected differentially by the Failure-Rest conditions.

As in the case of the I groups, the effects of the High and Low Motivation conditions were not significant in the analysis of variance of the adjusted recall scores under the delayed retention condition. The *F*-test for these effects in the D groups yielded a ratio of 1.70; whereas a value of 4.00 was necessary for significance at the 5% level.

The results of the present study, therefore, lend scant support to the contention that motivational differences will affect delayed recall even when they do not lead to differences in immediate recall. It will be recalled that Alper (1) suggested that under motivational conditions similar to those for the H groups in the present study there should be no significant drop in retention on the delayed test. However, as was noted above, all four delayed retention groups showed significant decrements in recall on the first relearning trial. It might at least be supposed, on the basis of Alper's report, that the H groups would nevertheless drop less than the L groups. The fact that the *F*-test for Low and High Motivation effects was not significant belies this expectation.

As was the case with the immediate

[5] This inference is, of course, contingent upon the use of an error term based upon the individual's variance as the best estimate of the population variance.

[6] See footnote 5, p. 212.

retention groups, none of the delayed retention groups differed significantly with respect to the number of trials required to reach the relearning criterion of two perfect trials.

Joint effects of failure, High and Low Motivation instructions and time elapsing before the relearning series.—The experimental design was set up so as to make possible statistical tests of any interaction among the variables. Accordingly, the two double-classification tables used to analyze immediate and delayed retention separately were combined to form a triple-classification table. The summary of the triple-classification analysis of variance is presented in Table 5.

Since the triple-interaction term proved nonsignificant, the simple interactions were tested, using the individual's variance as the error term. The interaction of failure effects with the time elapsing between original learning and relearning proved significant beyond the 1% level. This is consistent with the previous observation that the effect of failure varies with the time at which recall is measured. Failure effects influenced immediate recall markedly, but when a delayed measure of recall was taken, no influences attributable to failure were noted. It is this change in the effect of failure as a function of the time between original learning and the recall test which is reflected in the interaction term for these two variables.

It has been observed that the motivation variable showed no significant effects on the first relearning trial of either the groups tested immediately or of those tested 24 hr. later. The test for interaction between motivation and time elapsing between original learning and relearning correspondingly did not achieve significance. An F of 3.56 was obtained, whereas an F of 3.94 is necessary for significance at the 5% level for the *df* involved.

The interaction of the Failure-Rest variable with motivation did not prove significant in the triple-classification analysis. It will be recalled that neither of the double-classification analyses revealed any significant interaction between failure effects and the High and Low Motivation instructions. There is no reason, therefore, to suppose that the instructions given prior to original learning in this study had any differential effect upon the subsequent influence of the failure.

TABLE 5

TRIPLE-CLASSIFICATION ANALYSIS OF VARIANCE FOR ALL RECALL SCORES

Source	*df*	Sum of Squares	*F*
Motivation	1	.1953	
Failure	1	37.1953	
Time of Recall	1	321.9453	
M × F	1	4.1328	1.24
M × T	1	11.8828	3.56
F × T	1	29.0703	8.70*
M × F × T	1	.6328	
Individuals	120	400.9375	
Total	127	805.9922	

* Significant beyond the 1% level of confidence.

DISCUSSION

Three empirical issues concerning the effects of motivation and frustration upon one aspect of behavior, i.e., the retention of verbal material, provided the major impetus for this study. The relevance of the results for theoretical considerations of motivation and frustration in human learning and retention may be discussed in relation to these empirical problems.

The most striking result provided information relevant to the first question, which pertained to the effect of

verbally-induced failure in an experimental verbal learning task. It was clearly established that failed *Ss* showed a relative decrement in performance on the first trial of the test of retention taken immediately following the failure experience. No subsequent indices, either for relearning in the immediate retention groups or for all measures in the delayed retention groups, revealed any differences attributable to failure. Thus, the empirical picture shows disruption of performance for a short period of time as the only observed effect of failure in this situation. The hypothesis that failure would increase motivation, and that performance in failed groups would ultimately be superior to that in control groups as a result of this, was not supported. This is not crucial evidence, however, that failure experiences do not increase motivation. It may merely be that motivational effects, if they do exist, are not sensitively reflected in the type of task used here, where accuracy, rather than speed, was measured. Moreover, in learning situations where numerous "wrong" responses may be elicited, the effects of increased motivation are not easily predictable and may vary from *S* to *S*.

The ephemeral inhibitory effects of failure could be accounted for by considering the possible influence of any responses initiated by comments indicating failure. Such responses would be expected to interfere with performance to the extent that they proved to be incompatible with some of the complex of responses necessary for optimal performance on the recall task.

Just what mechanism might underlie the evocation by failure of responses which compete with those necessary for efficient performance on a given task is a matter of conjecture. Sears (6) has suggested that the experience of failure in card sorting results in a performance decrement by eliciting anticipatory goal responses which are incompatible with those conditioned to the instrumental act sequence ordinarily leading to successful achievement.

It is possible also that failure situations may produce a drive state (frustration, anxiety or insecurity) which affects behavior in this case through the elicitation of competing responses which have been associated with the accompanying drive stimuli.

It may be noted that observations of the depressive effect of failure have been largely limited to tasks in which verbal or symbolic processes play an important role. Farber, Russell, and Andreas (2), for instance, were unable to demonstrate that failure had any effect upon rotary pursuit performance. It may be that responses initiated by failure are verbal in nature. These verbal responses would be quite different from, and probably not incompatible with, the responses required for motor tasks. To the extent that the tasks used here required overt or implicit verbalization similar to that initiated by failure, interference with performance due to incompatibility of the responses would be expected.

These guesses as to the nature of the process leading to depressed performance under the conditions of the present study are not mutually exclusive. Several of the suggested factors, and, indeed, many others, may have operated to produce the failure effect.

If failure-induced responses underlie the immediate decrement in performance the temporary character of this decrement still remains to be explained. Apparently, any conflicting responses aroused by failure are rapidly extinguished. This could occur if the failure-produced responses were not strong enough to compete successfully with every "correct" response. The occurrence of any "successful" responses should result in the gradual reinstatement of the response complex necessary for efficient performance. Furthermore, since the failure comments in the present study were repeated only once, it would not be expected that responses to the failure situation would continue indefinitely. It may be that an occasional repetition of the failure experience would have maintained the conflicting responses and prolonged the period of depressed performance.

The second empirical question to be investigated was concerned with the influence of the initial motivating instructions upon the learning and retention of this task. The instructions used in this study informed the High Motivation *Ss* that the task was a test of scholastic aptitude. In this respect, the High Motivation *Ss* received instructions similar to those received by Alper's (1) "ego-oriented"

Ss, who were told that their task was an intelligence test. In addition, the use of a special incentive emphasized the importance of a good performance. The Low Motivation instructions stressed the routine, impersonal nature of the task even more than did Alper's "task-orientation" instructions. As far as instructions were concerned, then, the motivational differences between *Ss* of this study should have been at least as great as in the case of Alper's (1) study. To use her terminology, *Ss* in H groups should have been strongly "ego-involved," and those in the L groups only "task-involved."

Nevertheless, the results did not confirm Alper's contention that delayed recall under "ego-orienting" conditions would be different from delayed recall under "task-orienting" conditions. The adjusted recall score used in the analysis of performance on the first relearning trial was essentially the same as Alper's retention-difference score, upon the basis of which she concluded that there was no significant decrement in the recall of "ego-oriented" *Ss* after an interval of 24 hr. following original learning. With the methods used in this study, however, the High Motivation groups all showed a significant drop in performance on the delayed test. Furthermore, the decrement in recall for the High Motivation *Ss* did not differ substantially from that for the Low Motivation *Ss*. Thus, the adjusted recall scores offer no basis for the view that the motivational variable has any differential effect upon delayed recall when the level of performance in original learning is controlled.

So far as this present study is concerned, then, the only clear effect of the High Motivation instructions was the augmentation of the variability of performance during original learning. This result may indicate that reactions to instructions of this sort are a function of individual differences in personality characteristics. In a randomly selected sample from a normal population the number of individuals who improve under high motivating conditions and the number of individuals whose performance is impaired may balance, so that mean group performance is unaffected even though the variability within the group is increased. Until more empirical information is obtained concerning the effects of differentially motivating instructions more elaborate theoretical interpretations can scarcely be justified.

The third question dealt with the possibility that failure effects would vary as a function of *Ss* motivation level prior to the induction of failure. The results revealed no significant differences between the performance of High and Low Motivation *Ss* who were failed, and the interaction terms between failure and motivation were in all cases not significant. It can be concluded that failure effects in this type of learning task do not vary to any great extent as a function of the motivating instructions used. It remains a possibility that if the motivational differences were greater, or if some other aspect of behavior were measured, such a relationship might be found. This plausible hypothesis, nevertheless, still awaits confirmation.

Summary

1. Following differential instructions intended to produce high motivation in one-half of the *Ss* and low motivation in the remainder, 128 college students learned 12 nonsense syllables to the same criterion by the serial anticipation method. Half of the *Ss* in each motivation condition were then told that they were failing badly, while the others received no evaluation of their performance. Retention for the task was measured by relearning sessions held either 1 min. or 24 hr. after the original learning session. The retention conditions

were distributed equally within each of the groups previously defined. Thus, there were eight experimental groups with 16 *Ss* in each.

2. Relative to the performance of nonfailed *Ss*, failed *Ss* showed a highly significant disruption in performance on the first trial of the immediate test of retention. No subsequent measures, however, revealed any differences attributable to failure.

3. The variability of performance during original learning was significantly greater for those groups receiving the High Motivation instructions than for those receiving Low Motivation instructions. However, the mean number of trials required to reach the learning criterion was not significantly different from group to group. There was no reliable indication that the motivating instructions had an effect upon the amount of immediate or delayed retention or upon the magnitude of the failure effects.

4. Alper's contention that "ego-oriented" (highly motivated) *Ss* perform better than "task-oriented" (low motivated) *Ss* on delayed retention tests was not confirmed by the results of this study.

5. Consideration of performance in failed groups led to the conclusion that experimentally-induced failure had an inhibitory effect upon retention. This effect was transitory, however, dissipating both under conditions of rest and immediate relearning. A suggested explanation of this effect is offered in terms of competing verbal responses presumed to be aroused by failure induction. At least in this situation, no inference of motivational change following failure was necessary.

(Manuscript received July 16, 1951)

References

1. Alper, T. G. Task-orientation vs. ego-orientation in learning and retention. *Amer. J. Psychol.*, 1946, **59**, 236–248.
2. Farber, I. E., Russell, W. A., & Andreas, B. G. Effect of failure upon performance in verbal and motor learning situations. A review of work done at the State University of Iowa, reported to the Midwestern Psychological Association, April, 1949.
3. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.*, 1947, **18**, 50–55.
4. Melton, A. W. *Materials for use in experimental studies of the learning and retention of verbal habits.* Issued from the University of Missouri, 1940 (mimeo.).
5. Rosenzweig, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.
6. Sears, R. R. Initiation of the repression sequence by experienced failure. *J. exp. Psychol.*, 1937, **20**, 570–580.

DISPLACEMENT: GREATER GENERALIZATION OF APPROACH THAN AVOIDANCE IN A GENERALIZED APPROACH–AVOIDANCE CONFLICT[1]

NEAL E. MILLER AND DORIS KRAELING
Yale University

A previous paper has shown that the type of theoretical analysis that has been found useful in dealing with conflict behavior (2) can be extended to account for a considerable number of the details of the phenomena of displacement (3). This extension involves the assumption that the gradient of stimulus generalization of the inhibiting response is steeper (i.e., falls off more rapidly with stimulus dissimilarity) than that of the displaced response. This is illustrated in Fig. 1. It can be seen that in the original situation the inhibiting response is stronger than the displaced response so that the displaced response does not occur. In the new but somewhat similar situation, the strength of the displaced response is stronger than that of the inhibiting response so that the displaced response does occur. The assumption involved is analogous to the one that the gradient of avoidance falls off more steeply with distance in space than that of approach (2).

In a simple example, a man who is angry at his chief but inhibited from expressing any aggression against him takes it out on the office boy as a scapegoat. It is assumed that the aggression generalizes from the chief to the office boy. But we would also expect the responses inhibiting the aggression to generalize. And if the inhibiting responses generalize as strongly as the ones of aggression, the man should be just as unable to express aggression against the office boy as against his chief. Therefore, it is necessary to assume that the gradient of stimulus generalization of inhibition falls off more steeply than that of the responses that are inhibited.

The purpose of this study was to test the application of this assumption to a simple, controlled experimental situation in which an approach habit motivated by hunger is inhibited by an avoidance habit established by electric shocks.

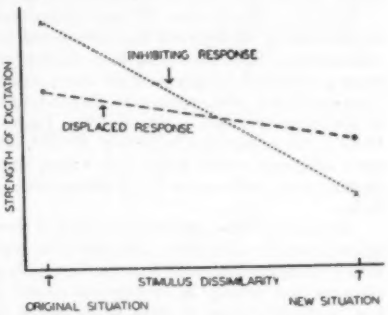

FIG. 1. Schematic representation of the assumption that the gradient of stimulus generalization for the inhibiting response is steeper than that for the displaced response. It is only for the sake of simplicity that the gradients are represented by straight lines in these diagrams; they could be any curves that have a continuous negative slope which is steeper for avoidance than for approach at each point above the abscissa.

[1] This study was supported by the Institute of Human Relations, Yale University. It was reported at the 1951 meetings of the Eastern Psychological Association. The authors thank Mr. G. K. Ogren for constructing the mechanical part of the apparatus and Dr. Martin Kohn for constructing the electrical part and for effective general assistance.

METHOD

General design.—First an approach habit was established by training hungry rats to run down a wide white alley to secure food. The food was concealed by a hinged shield which could easily be pushed out of the way. Then an avoidance habit, inhibiting the approach, was established by putting increasingly strong electric shocks on the food shield until the animals failed to touch it. Finally, the generalization of this approach-avoidance conflict was tested by comparing the response in the original stimulus situation with that in two new situations differing by varying amounts from the original one. This was accomplished by dividing the *S*s into three groups which were tested respectively (a) in the wide white alley (*same*), (b) in a narrower grey alley (*intermediate*), and (c) in a very narrow black alley (*different*).

As control for any effects of the alleys per se, half of the *S*s received their original training in the narrow black alley so that it was the *same* and the wide white alley was *different*. To control for the odor of the food, it was replaced on test trials by a pellet of moist clay. The possibility that *S*s placed in the different alley were not transfering specific habits, but only touching the pellet of clay by random activity was ruled out by a special control group. The control *S*s were kept on the same feeding cycle as the experimental ones, but instead of being trained to approach and avoid in the alley, they received pellets of food in restraining cages. Then they were tested in the same way as the experimental *S*s.

Subjects.—The *S*s were 115 male albino rats approximately 90 days old that completed the experiment. They were randomly divided into three groups: trained in wide white alley, trained in narrow black alley, and control *S*s not trained to approach or avoid in either alley. Each of these three groups was randomly divided into three subgroups tested in the wide white, intermediate grey, and narrow black alleys, respectively.

Apparatus.—The apparatus consisted of three separate alleys, each 36 in. long and 5 in. high with a transparent hinged lid, and a floor made of stainless steel bars $\frac{1}{16}$ in. in diameter spaced at $\frac{3}{4}$-in. intervals. One of these alleys was 6 in. wide with its walls and ends painted white, another was 4 in. wide with its walls and ends painted an intermediate grey, and a third was 3 in. wide with its walls and ends painted flat black.

At the starting end of each alley, the bars of the grid were connected to an electronic relay. As soon as the rat was placed into the alley, its feet shorted the grid allowing a subthreshold current to flow and actuate the electronic relay which started an electric timer.

FIG. 2. Diagram of the hinged shield which concealed the food. The round pellet of food rests on a metal strip inlaid into a bent strip of insulating plastic. When the rat touches the food, he completes a circuit from the metal strip to the grid below. The metal shield above normally hangs forward against the lip of the plastic so that the food is completely concealed. The shock can be connected to this shield and the grid.

At the far end of the alley was a little shelf on which food could be placed. This was covered by a hinged food shield which is illustrated in Fig. 2. The food shelf was connected so that when *S* touched a moist pellet of food, he completed a circuit to the grid below and thus caused a subthreshold current to operate an electronic relay which stopped the electric timer.

Above each food shield was a flashlight bulb that was brightly lighted for the white alley, lighted to an intermediate intensity for the grey one, and dimly lighted for the black alley. Unfortunately, the exact photometric specifications of these lights cannot be given. Each food shield was painted the color of its alley.

A fine grid of lines was cut through the paint so that when the shock device was connected *S* would receive a shock whenever he touched the food shield with his nose and completed a circuit to the grid. The amount of current used as a shock was kept relatively constant in spite of fairly wide fluctations in *S*s' resistances by the pentode tube circuit described by Muenzinger and Waltz (5).

Feeding schedule.—The *S*s were fed dry lab chow biscuits at the same hour each day shortly after the end of training. In order to keep them on a reduced diet, the water bottles were re-

moved from their cages while they were left with dry food for 1 hr. After this the food was removed and the water restored. During taming and training the reward consisted of small pellets, approximately the size of a pea, made of moist, ground lab chow.

Taming.—On the second day of reduced feeding *S*s were tamed by receiving three trials with the pellet of wet food on a board in the bottom of individual cages and three trials with food on a similar board outside the cages with each *S* held by hand. On the fifth day of reduced feeding, they were given six more trials eating food from the board while being held by hand.

Training.—Training to approach started on the sixth day of reduced feeding. It consisted of 50 trials distributed over a period of ten days in the following way: 3, 5, 5, 5, 5, 0, 0, 10, 10, 10, 2. Animals were run in rotation in groups of six so that at least 60 sec. intervened between successive trials. Each *S* was placed at the start of the alley and the time for him to run down, push back the food shield, and touch the food was recorded. On the first two trials of the first day, the food shield was propped back so that the food was clearly visible; after that, *S* had to push the shield back to get the food. Animals that did not eat within 5 min. on any of the first four trials, within 1 min. from then through the fourth day of training, or within 5 sec. on any two consecutive trials thereafter were discarded. Six *S*s were discarded for these reasons.

Feeding in the home cages was omitted on the next to the last day of training so that *S*s were 48-hrs. hungry on the day when they received the last two trials of approach training, the avoidance training, and the test trials.

Avoidance training started immediately after the last two approach trials. On each avoidance trial *S* was placed in the alley for 60 sec. with the shield connected so that he received an electric shock whenever he touched the shield with his nose or paws. The shock was .1 ma. on the first trial and was increased by .05 ma. after each trial on which *S* received a shock. Avoidance training was continued until *S* failed to receive shock on three successive 60-sec. trials.

The control *S*s were tamed in the same way as the experimental ones, kept on the same feeding cycle, but were fed pellets of food in special cages instead of being given approach and avoidance training in the alley. On the last day of training, they were randomly paired with experimental *S*s. Half of them preceded their partner and half of them followed him. Thus they received their pellets in the same sequence as the experimental *S*s.

Test trials.—The test trials began immediately after *S* had reached his criterion of three successive trials without shock. A control *S*'s test trials started after his partner had reached the criterion. On test trials each of the three groups was randomly divided into three subgroups, one of which was tested in each of the three alleys. The goal end of the alley was carefully wiped out to remove traces of food and a little pellet of moist clay was inserted in place of the usual one of food. The time to run down, push aside the shield, and touch the pellet of moist clay was automatically recorded by the electronic relay and clock. Animals not running within 60 sec. were removed to wait their turn for another trial up to a maximum of four. During test trials *S*s were run in rotation so that an average of approximately 5 min. elapsed between successive trials for each *S*.

Results and Discussion

The percentage of animals reaching the goal during the four test trials is shown in Fig. 3. (An analysis in terms of the total time to reach the goal on test trials yielded exactly the same pattern of results but was somewhat less reliable statistically.) When the test alley was the same as the training one, only 23% of the *S*s reached the goal during the test trials. In the *intermediate* and *different* alleys progressively increasing percentages, 37% and 70%, reached the goal. This is exactly what would be expected from the assumption that the avoidance generalizes less strongly to the new situation than the approach. The difference between the groups tested

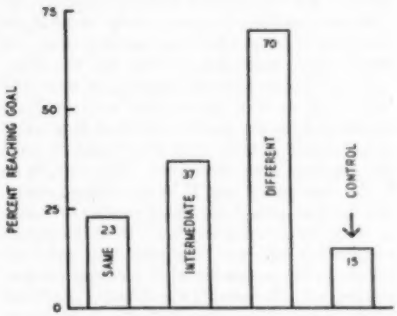

Fig. 3. The percentage of *S*s in each group reaching the goal during the four test trials

in the *same* and *different* conditions is significant at the .001 level of confidence.[2]

That the success of the *Ss* tested in the *intermediate* and *different* alleys is due to generalization from previous training rather than to mere random behavior without any specific transfer is shown by the fact that only 15% of the control *Ss*, which did not receive any previous training, reached the goal. The difference between the control group and the one tested in the different alley is significant at the .001 level. Furthermore, *Ss* tested in the two new stimulus situations showed definite signs of conflict, with increasingly hesitant approach in the neighborhood of the food shield punctuated by sudden withdrawals.

Thus, it may be concluded that the approach-avoidance conflict generalized to the new stimulus alleys, and that the avoidance generalized less strongly than the approach.

In this experiment, the approach habit was established first and the avoidance one later. A subsequent experiment (6) using the strength-of-pull technique to measure the two habits separately in different groups of *Ss*, shows that the gradient of stimulus generalization is steeper for avoidance than for approach when the factors of age of habit and distribution of training trials are controlled.

As has been shown elsewhere (2, 3, 4), the assumption confirmed in these simple, controlled experimental situations is useful in extending a theoretical model of conflict behavior to phenomena of displacement and psychotherapy.

Summary

1. This experiment was designed to test a simple deduction from the assumption that the gradient of stimulus generalization of avoidance falls off more steeply than that of approach.

2. An approach-avoidance conflict was established by training hungry albino rats to run down a short alley for food and then giving them electric shocks till they failed to touch the food on three successive trials.

3. To study the generalization of this conflict to new stimulus situations, some *Ss* were tested in the same alley in which they had been trained, others were tested in a somewhat different alley, and still others in a quite different alley. In order to determine the effects of mere random behavior in a new situation, a control group was tested without previous training in any alley.

4. The experimental *Ss* tested in the different alleys showed distinctive vacillation, with increasingly hesitant approach punctuated by sudden withdrawals, clearly indicating that they had generalized the conflict to the new situation.

5. During the four test trials the proportion of *Ss* approaching near enough to touch the "food" was 15%

[2] Although all *Ss* were brought up to the same criterion of avoidance (and hence supposedly had the same balance of conflict), it was found that the number of shocks required to bring them to this criterion averaged 4.4, 4.2, and 3.0 for *Ss* to be tested, respectively, in the *same*, *intermediate*, and *different* alleys. In spite of the fact that *Ss* had been selected randomly, the difference between the groups to be tested in the *same* and *different* alleys was reliable at the .05 level. Since inspection showed that the difference was due to a few *Ss* requiring an unusually large number of shocks in the *same* group, a special analysis was made in which all *Ss* in either group requiring more than five shocks to reach the criterion were eliminated. This left 16 *Ss* in the *same* group, and 21 in the *different* group, and for both groups the mean number of shocks to reach the criterion was 2.5. With the groups equalized in this way, the conclusion remained the same: the proportion of *Ss* reaching the goal was larger for *Ss* tested in the different alley than for those tested in the same alley. A chi-square test corrected for discontinuity showed that the difference was significant at the .003 level.

for the controls, and for the experimental animals 23% in the same alley, 37% in the somewhat different alley, and 70% in the quite different one. This shows that avoidance generalized less strongly than approach to the new stimulus situations.

(Manuscript received July 24, 1951)

References

1. Dollard, J., & Miller, N. E. *Personality and psychotherapy: An analysis in terms of learning, thinking, and culture.* New York: McGraw-Hill, 1950.
2. Miller, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944.
3. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
4. Miller, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. *J. Pers.*, 1951, **20**, 82–100.
5. Muenzinger, K. F., & Waltz, F. C. An analysis of the electrical stimulus producing a shock. *J. comp. Psychol.*, 1932, **13**, 157–171.
6. Murray, E. J., & Miller, N. E. Displacement: Steeper gradient of generalization of avoidance than of approach with age of habit controlled. *J. exp. Psychol.*, 1952, **43**, 222–226.

DISPLACEMENT: STEEPER GRADIENT OF GENERALIZATION OF AVOIDANCE THAN OF APPROACH WITH AGE OF HABIT CONTROLLED [1]

EDWARD J. MURRAY AND NEAL E. MILLER
Yale University

Miller (3) has shown that the assumptions employed in the analysis of conflict behavior (2) may be fruitfully extended to displacement. One of the key assumptions is that the gradient of stimulus generalization of avoidance is steeper than that of approach. Miller and Kraeling (4) have secured evidence supporting this assumption. They have shown that when an approach-avoidance conflict is generalized to a new stimulus situation, rats are more likely to resume approaching the dangerous goal. From this it can be inferred that the strength of the avoidance motivated by fear must have been reduced more than that of the approach motivated by hunger. But their procedure necessarily involved teaching the animals to approach the goal first, before avoidance was established by giving them electric shocks at the goal. Thus, the approach was an older habit than the avoidance. Moreover, the approach was established in a large number of distributed trials while the avoidance was learned in a small number of relatively massed trials.

The purpose of the present study was to control for the age of the habit and the number and distribution of training trials by making separate measures of generalization on two groups of *S*s trained to approach and avoid, respectively, with the same schedule of training trials. A subsidiary purpose was to see whether the relationship between the two gradients would remain the same if the procedure of partial reinforcement was used.

Procedure

Experiment with 100% Reinforcement

Subjects.—The *S*s were 15 male albino rats of Wistar strain approximately 90 days old at the beginning of the experiment. They ranged in weight from 264 to 340 g. Two groups of 7 and 8 were matched on the basis of weight.

Apparatus.—The apparatus consisted of two experimental alleys, a shocking device, and a setup for measuring strength of pull.

The two alleys differed markedly in the following respects: One was 6 in. wide, painted white, and placed directly under the overhead light in the room. This will be called the *white alley*. The other was only 3 in. wide, painted black, and placed at a distance from the overhead light. This will be called the *black alley*.

Each alley was 10 in. deep, and each floor was a grid of stainless steel bars $\frac{1}{16}$ in. in diameter spaced at $\frac{1}{2}$-in. intervals. The first 8 in. of each alley was a starting compartment separated from the main part of the alley by a metal door which *E* operated by hand. Opening the door automatically turned on the current during shock trials. The main part of the alley was 32 in. long.

In order to allow *S*s motivated by electric shock to escape it by running down the alley, each alley contained an "island of safety" 12 in. long, 1 in. high, and insulated from the grid. A small food cup was placed on the far end of this island. This "island of safety" was movable and could be placed 0, 6, 13, or 20 in. away from the starting-box door. Wherever it was placed a wooden panel was inserted in the alley just behind it so that the far end of the "island of safety" was the end of the alley.

[1] This study was supported by the Institute of Human Relations. It was presented as a paper at the Eastern Psychological Association meetings in 1951. The authors wish to express their gratitude to Mr. K. G. Ogren for help in constructing the apparatus.

The shock used in this experiment was 310 v. of 60-cycle AC, after potentiometer adjustment of voltage supplied by a radio transformer, and administered through a 250,000-ohm resistance in series with *S*.

Strength of pull was measured by a modification of the apparatus used by Brown (1). A simple rubber-band harness was placed with one loop in front of the rat's forelegs and one loop behind them. A cord was run back from the harness to a metal block sliding on a rod. This metal block was connected to an elastic strip of rubber. As the rats pulled the block along the rod against the tension of the rubber strip, a marker attached to the block made a line on waxed polygraph paper. From a calibration with weights, the length of this line could be converted to the strength of pull in grams (e.g., 1 in. = 83 gm.). The slack in the cord was such that the rat took it up and began pulling against the marker when his nose was 10 in. beyond the starting-box door.

Taming.—After 48 hr. of water deprivation *Ss* were allowed to drink from a bottle while being held by hand. This was repeated on the next day, after which water was permanently restored to the living cages. Then *Ss* were put on a 24-hr. feeding rhythm for five days. During the first two days they were fed wet mash while being handled. Thereafter they were fed ordinary Purina dog chow checkers for 1 hr. at the same time each day. This feeding schedule and the value of the shock were selected on the basis of exploratory work to give about the same level of performance on test trials.

Training.—Eight *Ss* were trained with food as the reinforcement and seven with escape from shock as the reinforcement. The training of the shock and the hunger *Ss* was made as comparable as possible. The hunger drive was the same for both groups; 1 hr. after their day's training all *Ss* were fed for 60 min.[2] The reward for the hunger *Ss* was a pellet about the size of a pea made of ground Purina chow, moistened with water and enriched with a little cod liver oil and yeast. In order to keep the shock *Ss* on exactly the same diet, they were given the same number of these pellets during their feeding hour.

On each trial the hunger *Ss* were placed gently in the starting box and the starting door was raised. They ran down the alley, climbed up on the "island of safety," and picked up the food pellet in the cup. After nearly finishing the pellet, they were picked up gently and put back in the restraining cage. The shock *Ss* were handled in the same way. The *E* gently placed them in the starting compartment and raised the door which turned on the shock. About 10 sec. after *S* reached the "island of safety," he was gently removed from the apparatus.

In order to control for the effect of the alley itself, half of the *Ss* in each group were trained in the white alley and half in the black one. They were run in rotation in groups of four: white-shock, white-hunger, black-shock, and black-hunger. Each *S* was given four trials a day with about 5 min. separating each trial. Between trials they were kept in individual restraining cages.

Five such days of training were given, making a total of 20 trials. In order to facilitate learning, the first 6 trials for both the hunger and the shock *Ss* were given with the "island of safety," which also contained the food cup, 0 in. from the starting compartment door, the next 4 at 6 in., then 2 at 13 in., and finally, the last 8 at 20 in. During the 4 trials of the last day, the animals wore the rubber-band harness.

Testing.—In order to reduce the amount of learning occurring to the new stimulus context during tests for generalization, the test trials were nonreinforced. The *Ss* trained with hunger were tested without food present; *Ss* trained with shock were tested without it. It should be noted that this duplicates the conditions of a conflict in which *S* remains some distance away from the feared goal and hence does not receive either food or shock. Since *Ss* trained with shock were motivated only by fear on test trials, they will be referred to hereafter as the *fear* group.

Each *S* was given two tests in the *same* alley he had been trained in and two in the *different* alley. Any sequential effects were controlled by testing half of each subgroup in an SDSD order and the other half in a DSDS order.

During each test *S* was restrained halfway to the goal for 5 sec. and then released. A record was made of the maximum distance that *S* pulled the marker against the tension of the elastic strip of rubber.

Experiment with Partial Reinforcement

In this experiment the *Ss* (15 new rats ranging in weight from 204–380 gm.), the apparatus, and the procedure were the same as in the previous one with the following exception: The factor of partial reinforcement was introduced by adding seven nonreinforced trials during the last two days of training. On these trials the shock was omitted for *Ss* rewarded by escape from shock,

[2] The procedure of having the avoidance *Ss* hungry during training and testing also helped to make the conditions similar to those encountered in experiments on conflict (2, 4). In such experiments the *Ss* are hungry as well as frightened.

and the food was omitted for *S*s rewarded by eating food. In order to maximize the effect (6), each nonreinforced trial was given immediately before a reinforced one.[3] The interval following each reinforced trial remained the same, approximately 5 min. The nonreinforced trials were given immediately before the second, third, and fourth regular reinforced trials on the fourth day and immediately before all four trials on the fifth day.

Results and Discussion

The results of the four test trials given to the group with 100% reinforcement are shown in Fig. 1. It can be seen that the gradient of generalization of the avoidance response motivated by fear is steeper than that of the approach response motivated by hunger. Since the two gradients cross, the difference in slope could not have been an artifact of distortions in the units of measurement. The results of all groups are presented in more detail in Table 1, which shows that the generalization decrement of the fear *S*s is highly reliable whereas that of the hunger *S*s is not. The difference between the decrements (i.e., steeper slope of avoidance) is highly reliable.

When only the first test trial is used, the gradient of generalization of fear is much flatter. Neither of the decrements is reliable nor is the difference between them. These results are in line with the general finding that there may be little difference between the original and generalized response on the first test trial but that the latter have much less resistance to extinction by nonreinforced trials (1).[4]

It is also quite possible that the strength of avoidance motivated by fear on the first test trial in the original situation is spuriously low, being reduced by factors of the type that N. E. Miller and D. Lawrence have observed in an unpublished study. They have observed that *S*s, trained to run to escape electric shock, slow down markedly immediately after the shock is turned off, tending to crouch, but that during successive *shockless* trials these *S*s progressively *increase* their speed. This observation shows that the response on the first shockless trial does not indicate the full strength of potential response. Their observations suggest (a) that the response of running does not generalize completely from the drive of pain to that of fear, (b) that the fear elicits an innate response of crouching which tends to be incompatible with running, and (c) that running is relearned (and crouching extinguished) as a response to fear during the first few shockless trials. That something of this kind may have occurred in the present experiment is suggested by the fact that the fear *S*s showed a progressive increase in response in later tests in the original situation while the general trend of the fear-different group and both groups of hunger *S*s was downward. One of the theoretical reasons for introducing some partial reinforcement into the second experiment was to give the avoidance *S*s some shockless training trials during which running would be reinforced as a response to fear-without-shock while crouching would be extinguished. This should tend to remove any spurious reduction in the strength of pull in the first test trial in the original avoidance situation and hence produce a steeper gradient.

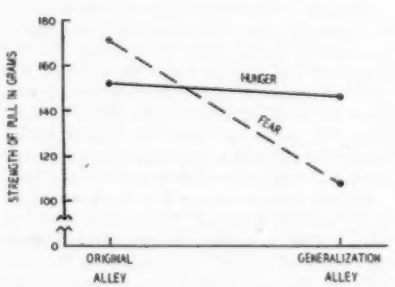

Fig. 1. Gradients of stimulus generalization of avoidance motivated by fear and approach motivated by hunger. The strength of pull of each *S* was measured twice in the same alley in which he was originally trained and twice in a different alley.

[3] While primary reinforcement was absent on these trials, a considerable amount of secondary reinforcement probably was present, mediated by cues from the "island of safety."

[4] Such extinction, increasing the slopes of the gradients, could occur during the course of a conflict trial.

TABLE 1

AVERAGE STRENGTH OF PULL IN GRAMS ON TEST TRIALS

Test Conditions	100% Reinforcement				Partial Reinforcement			
	Trial 1		Av. 2 Trials		Trial 1		Av. 2 Trials	
	Mean	p*	Mean	p*	Mean	p*	Mean	p*
1. Fear-original	146.8	.45	171.1	.001	155.7	.02	158.6	.003
2. Fear-different	138.0		108.5		124.2		110.7	
3. Hunger-original	174.2	.35	152.3	.30	128.5	.35	117.9	.45
4. Hunger-different	161.8		145.1		142.5		120.8	
5. (1–2) – (3–4)		.49		.002		.07		.004

* The p-values are from t-tests for evaluating differences in the predicted direction and are therefore based on one tail of the probability distribution.

As Table 1 shows, on the first test trial the fear group trained with partial reinforcement showed a generalization decrement which is statistically reliable. The hunger group showed a slight difference in the opposite direction which is not reliable. The greater steepness of the gradient of fear (i.e., algebraic difference between the differences) approaches reliability on the first test trial and becomes highly reliable when all four test trials of the Ss trained with partial reinforcement are used.

We may conclude that when the age of habit and number of training trials are held constant, the generalization gradient of avoidance motivated by fear is steeper than that of approach motivated by hunger. When Ss are trained with 100% reinforcement, this difference does not show up on the first test trial. When they are trained with partial reinforcement, it does. With either kind of training it shows up clearly in a series of nonreinforced trials.

In these experiments the approach was motivated by hunger and the avoidance by fear. One might attempt to explain the difference in terms of appetitive and aversive drives. However, according to an hypothesis put forth by Miller (**2, 3**) and Brown (**1**), the difference is due to the fact that the approach is based on a primary drive and the avoidance on a learned, response-produced one. This hypothesis has been confirmed in an experiment (**5**) in which the gradient of a response motivated by the learned drive of fear was found to be steeper than that of a response motivated by the primary drive of pain.

SUMMARY AND CONCLUSIONS

1. The purpose of this study was to compare generalization gradients of avoidance motivated by fear and approach motivated by hunger when the age of the habit and the number and distribution of training trials were controlled.

2. Two groups of rats were trained to run down an alley. One received food as reinforcement and the other escape from shock as reinforcement. Another pair of groups was trained in the same way except that seven nonreinforced trials (i.e., partial reinforcement) were added to the training

procedure. All *S*s were given four nonreinforced test trials on which their strength of pull was measured. No shock or food was present on these trials. Two test trials were in the *same* alley in which they were trained and two in a *different* alley. The sequence of testing was balanced.

3. The difference between the strength of pull in the same alley and in the different one (i.e., steepness of the gradient of stimulus generalization) was greater for the avoidance response motivated by fear than for the approach response motivated by hunger. For *S*s trained with 100% reinforcement this difference did not show up on the first test trial; for those trained with partial reinforcement, it did. For both the 100%- and the partial-reinforcement groups this difference in slope was highly reliable when the scores of the nonreinforced test trials were averaged.

(Manuscript received July 24, 1951)

References

1. Brown, J. S. Gradients of approach and avoidance responses and their relation to level of motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.
2. Miller, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944.
3. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
4. Miller, N. E., & Kraeling, D. Displacement: Greater generalization of approach than avoidance in a generalized approach-avoidance conflict. *J. exp. Psychol.*, 1952, **43**, 217–221.
5. Miller, N. E., & Murray, E. J. Displacement and conflict: Learnable drive as a basis for the steeper gradient of avoidance than of approach. *J. exp. Psychol.*, 1952, **43**, 227–231.
6. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–526.

DISPLACEMENT AND CONFLICT: LEARNABLE DRIVE AS A BASIS FOR THE STEEPER GRADIENT OF AVOIDANCE THAN OF APPROACH[1]

NEAL E. MILLER AND EDWARD J. MURRAY

Yale University

A key assumption in conflict theory is that distance of *S* from the feared goal produces a greater decrement in the strength of responses of avoidance than it does in responses of approach (6, 8, 10). In short, the gradient of avoidance is assumed to be steeper than that of approach. The same type of theoretical analysis can be extended to include phenomena of displacement (11) and psychotherapy (5) by making the similar assumption that the gradient of stimulus generalization of avoidance is steeper than that of approach. The functional similarities between the gradients in space and those in stimulus generalization are in line with the hypothesis that the former are derived from the latter (2, 3, 5, 7, 17).

The assumption that the gradient of avoidance is steeper in both distance and stimulus dissimilarity than that of approach is confirmed by experimental and clinical evidence from a variety of situations (3, 5, 10, 11, 15, 16).[2] But such evidence does not prove that this assumption is applicable to all situations. An analysis of the reasons for the greater steepness of the gradient of avoidance should help us to define more exactly the generality of this phenomenon.

A previous study has shown that the difference in steepness persists when the factors of age of habit and number and distribution of training trials are controlled (16).

Miller (10, 11) and Brown (3) have advanced the hypothesis that the difference in slope is produced by a difference in the source of the drive motivating the behavior. In all the experiments the approach has been motivated by the primary drive of hunger while avoidance was motivated by the learned drive of fear. According to a theory of learned drive (13, 14), such drives are produced by internal responses which obey the same laws as overt ones. Thus, when the avoidance habit is tested in a new stimulus situation, both the fear motivating the avoidance and the overt responses of avoidance should be weakened by generalization. On the other hand, the hunger motivating the approach is a primary drive more dependent upon physiological factors which the subject carries with him into new stimulus situations. Thus generalization will weaken *only* the responses of approach; it will *not* also weaken the drive motivating the responses. With the drive remaining more constant, the gradient of approach should be less steep than that of avoidance.

The preceding analysis has dealt with the effects of stimulus dissimilarity on the gradient of stimulus generalization; a parallel analysis can be made for the effects of spatial distance on the gradient of delay in reinforcement.

[1] This study was supported by funds from the Institute of Human Relations, Yale University. The authors thank Mr. K. G. Ogren for help in constructing the apparatus.

[2] It should also be noted that alcohol and the barbiturates seem to produce a greater reduction in the strength of avoidance than of approach (1, 4, 5, 9).

The foregoing theoretical analysis leads to the prediction that the gradient of stimulus generalization of an avoidance habit motivated by the learned drive of fear should be steeper than that of an avoidance habit motivated by the primary drive of mild pain produced by an electric shock.[3] The purpose of the present experiment was to test this prediction.

Method

General design.—First, albino rats were trained to escape electric shock by running down a wide white alley, the floor of which was a grid, to a little elevated "island of safety." Then half of them were tested with the shock turned off so that they were motivated by the learned drive of fear without any primary drive of pain. These *S*s will be called the *fear* group. The other half were tested with the shock on so that they were motivated by the primary drive of pain, and perhaps by some fear. These animals will be called the *pain* group.

In order to test for stimulus generalization, half of each group was tested in the original wide white alley while the other half was tested in a different narrow black one. The effects of the alleys per se were balanced out by training an equal number of additional *S*s in the narrow black alley so that in the test it was the same and the wide white one was different.

The strength of response was measured by having *S*s wear a little harness attached to a cord, stopping them halfway down the alley for 5 sec. on the test trial, and recording their maximum strength of pull during this period. But it is possible that there is a curvilinear relationship between physically equal units of pull in grams and psychologically equal units of response strength. For example, as *S*s approach their limit of strength, large increments in excitatory potential might be required to produce small increments in grams of pull. Thus a gradient measured at this high level would have its apparent slope spuriously reduced. In order to control for this possibility, the fear and pain groups were tested at the same general level of performance. This gave the gradients a chance to cross if there was any difference in their slopes. This equalization of general level of test response was accomplished by giving the fear group (tested without shock) its original training with a somewhat stronger shock than the pain group (tested with shock).

In order to control for the effects of this stronger shock used during the training of the fear group, an additional pain group was run. This group was trained with the same stronger shock as the fear group and also tested with that shock. In order to prove that any flat gradient exhibited by these *S*s was not due to having reached their limit of strength, they were given an additional test with a stronger shock.

Subjects.—The *S*s were 60 male albino rats of Wistar strain approximately 90 days old at the beginning of the experiment. Of these, 45 were used in the main experiment and 15 as the special group trained and tested with the stronger shock. Four additional *S*s were started but not tested because they failed to learn.

Apparatus.—The apparatus is described in more detail elsewhere (16). It consisted of one wide (6 in.) white alley, and narrow (3 in.) black alley, each 40 in. long and with a grid for a floor. The first 8 in. was a starting compartment. Opening the door out of this compartment turned on the electric shock for the whole grid. A movable "island of safety," 1 in. high and 12 in. long, allowed *S*s to get off the grid and escape shock.

On test trials the strength of pull was recorded by a rubber-band harness connected to a cord. When *S* had run halfway down the alley, the cord engaged a marker and pulled it against the force of a strip of rubber. At the end of 5 sec. of restraint *E* pulled a simple catch releasing *S*.

Training.—During the first seven days *S*s were tamed and a feeding rhythm established by a procedure that has already been described (16). In order to keep the conditions of this experiment parallel with those of preceding ones on conflict and displacement (10, 15, 16), *S*s were kept hungry by 22 hr. of food deprivation with 1 hr. of free access to dry food which began 1 hr. after the end of each day's trials.

Training consisted of four trials a day for five consecutive days. The *S*s were run in rotation in groups of four with a period of at least 5 min. between consecutive trials.

On each training trial *S* was placed in the starting box. Approximately 5 sec. later the door was raised, automatically turning on the shock. Approximately 10 sec. after *S* reached the "island of safety" he was removed and placed in a restrain-

[3] If it is possible to establish a strong enough learned drive (appetite) so that *S*s will continue to approach when hunger is satiated, a similar prediction can also be made for approach habits. The gradient of generalization of an approach habit should be steeper when it is motivated only by the learned components of appetite than when it is motivated by innate components of hunger. The authors plan to try to test this prediction in a subsequent experiment.

ing cage. In order to facilitate learning, the "island of safety" was placed up against the starting-box door for the first six trials. After this it was progressively moved back with 4 trials at 6 in., 2 at 13 in., and 8 at 20 in. During the last four trials of training, *Ss* wore the rubber-band harness.

For the fear group (trained with shock, tested without shock) the shock was 310 v. of 60-cycle AC administered through a 250,000 ohm resistance in series with *S*. For the pain group (trained and tested with shock) it was 100 v. The special strong-shock, pain group received 310 v. during training and on the first test; the additional test was with 410 v. Half of each group was trained in the white alley and half in the black one, making a total of six training subgroups.

Testing.—Each *S* was tested on the day after the end of training. Half of the *Ss* in each of the training subgroups were tested in the same alley that they had been trained in, the other half were tested in the other alley. During test trials the rubber-band harness was connected to a cord leading to the polygraph measuring the strength of pull. When they were halfway to the "island of safety," this cord engaged the marker on the polygraph so that they pulled against the elastic strip of rubber for 5 sec. The fear group were tested without shock; the pain group was tested with the usual 100-v. shock. The special strong-shock, pain group was tested with a 310-v. shock and then given an additional test with 410 v.

Results and Discussion

The results are presented in Fig. 1. It can be seen that the gradient of stimulus generalization is steeper for the group motivated by the learned, response-produced drive of fear than for the one motivated by the primary drive of pain. Because the two gradients cross, the difference in slope could not have been an artifact of any distorted relationship between physical units of pull and psychological units of response strength. In order to test for reliability, an analysis of variance was performed. In this analysis the difference in slope showed up as the interaction term. The *F* of this term was 6.7 for 1 and 41 *df*; it is significant in the predicted direction beyond the .01 level.

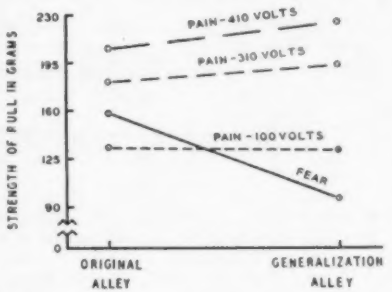

FIG. 1. The generalization of responses motivated by the learned drive of fear and by the primary drive of pain

This result is what was predicted from the theory of learned, response-produced drive. It shows that the difference between an aversive and an appetitive drive is not a necessary condition for securing the difference in slope.

That the difference in slope was not due to the fact that the fear group received stronger shocks during training is suggested by the fact that the special group trained and tested with this stronger shock (310 v.) did not show a generalization decrement. Indeed there is a slight reversal which is not significant ($p = .20$, using one tail). An analysis of variance between this high pain group and the fear group shows that the difference between the slopes is highly reliable ($p = .003$).

That this lack of a difference between the high-pain groups tested in the same and different situations was not due to the fact that *Ss* were pulling as hard as they could in both of them is shown by the fact that an increase in the strength of the shock (to 410 v.) produced an increase in the strength of pull. Averaging the same and different tests, this increase is highly reliable ($p = .001$, using one tail).

It is logically possible to assume that a stronger shock during training produces a steeper gradient of general-

ization but that, even though *S*s in the special group trained and tested with strong shock were not at their physiological limit, the function relating excitatory potential to strength of pull is just sufficiently negatively accelerated to mask the difference in slope produced by the stronger shock. This type of explanation is so cumbersome, however, that it does not seem very probable.

In this experiment the primary drive was actually dependent upon an external source of electric current, but this was kept constant in the different test situations, much as the primary physiological component of hunger would be expected to remain constant in different stimulus situations. It is entirely possible that *S*s in the pain group were also responding with a learned drive of fear. Any such fear as a part of the motivation of the pain group would be expected to *reduce* the difference between the groups.

Finally, it should be noted that the results of this experiment, indicating that a difference in the type of drive can produce a difference in the slope of the gradient of generalization, do not rule out the possibility that other factors may also affect the steepness of the gradients. For example, one would expect discrimination training to be able to increase the steepness of either of the gradients (**10, 11**).

Summary

1. The purpose of this experiment was to test the prediction that the gradient of stimulus generalization should be steeper for responses motivated by a learned, response-produced drive dependent upon external cues than for a primary drive independent of external cues.

2. Albino rats were trained to escape electric shock by running down a distinctive alley. Then half of them were tested with the shock turned off so that they were motivated only by the learned drive of fear which would be expected to vary with the cues in the alley. The other half were tested with the shock on so that they were motivated by the primary drive of pain which would not be expected to vary with cues such as the color and width of the alley. In order to test for stimulus generalization, half of the fear and half of the pain group were tested in the same alley in which they were trained; the other halves were tested in a different alley. During these tests the strength of the response was measured by the force with which the animals pulled against a temporary restraint.

3. The gradient of stimulus generalization of the group tested with the learned, response-produced drive of fear was steeper than that of the group tested with the primary drive of pain. This result shows that the difference in slope is not produced by the difference between appetitive and aversive drives, and confirms the theoretical prediction.

4. The fact that the two gradients crossed rules out any artifact that might be produced by a distorted relationship between physical units of pull and psychological units of response strength. Other controls make alternative explanations seem quite improbable, but cannot entirely eliminate them.

(Manuscript received July 24, 1951)

References

1. Bailey, C. J., & Miller, N. E. Effect of sodium amytal on behavior of cats in an approach-avoidance conflict. *J. comp. physiol. Psychol.*, 1952, **45**. (In press.)
2. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33**, 209–226.

3. Brown, J. S. Gradients of approach and avoidance responses and their relation to level of motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.

4. Conger, J. J. The effects of alcohol on conflict behavior in the albino rat. *Quart. J. Stud. Alcohol.*, 1951, **12**, 1–29.

5. Dollard, J., & Miller, N. E. *Personality and psychotherapy: An analysis in terms of learning, thinking, and culture.* New York: McGraw-Hill, 1950.

6. Hull, C. L. The goal-gradient hypothesis applied to some "field-force" problems in the behavior of young children. *Psychol. Rev.*, 1938, **45**, 271–299.

7. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.

8. Lewin, K. Environmental forces in child behavior and development. In C. Murchison (Ed.), *A handbook of child psychology.* Worcester: Clark Univ. Press, 1931.

9. Masserman, J. H. *Behavior and neurosis.* Chicago: Univ. Chicago Press, 1943.

10. Miller, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944.

11. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.

12. Miller, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. *J. Pers.*, 1951, **20**, 82–100.

13. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

14. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.

15. Miller, N. E., & Kraeling, D. Displacement: Greater generalization of approach than avoidance in a generalized approach-avoidance conflict. *J. exp. Psychol.*, 1952, **43**, 217–221.

16. Murray, E. J., & Miller, N. E. Displacement: Steeper gradient of generalization of avoidance than of approach with age of habit controlled. *J. exp. Psychol.*, 1952, **43**, 222–226.

17. Spence, K. W. The role of secondary reinforcement in delayed reward learning. *Psychol. Rev.*, 1947, **54**, 1–8.

AN EQUAL DISCRIMINABILITY SCALE FOR LOUDNESS JUDGMENTS [1]

W. R. GARNER

The Johns Hopkins University

In a previous paper, Garner and Hake (3) described a technique for constructing a scale of equal discriminability with data obtained from absolute judgments of a stimulus magnitude, or stimulus ratings. Basically, *S* is required to make judgments about single stimuli by assigning each stimulus to one of several response categories. Then a scale of discriminability for the several response categories is constructed from the overlaps of the judgmental dispersions, and the median value for each stimulus on this discriminability scale is determined. The technique produces a scale in which equal distances on the scale represent equal tendencies to judge two stimuli as being in the same category. Thus the scale is truly one of discriminability of the stimuli, but discriminability for absolute judgments rather than for comparative judgments. The technique proposed leads to the same type of scale as techniques previously proposed by Saffir (6) and Attneave (2).

Preliminary unpublished data by the author on such an equal discriminability scale for loudness suggested several hypotheses about the nature of such a scale and the experimental factors which affect it. The present experiments were designed to test specifically three hypotheses.

1. The equal discriminability scale for loudness is linearly related to a scale of cumulative difference limens for intensity. Thus the discriminative process is the same whether comparative or absolute judgments are being made.

2. In an experimental series, the fact that the series has definite limits produces anchoring effects which make discriminability relatively better near the limits. Thus discriminability will be better near the highest and the lowest intensities presented in the series.

3. The particular spacing of the stimulus intensities presented has no effect on the discriminability scale as long as the limiting intensities are the same. Thus the shape of the discriminability scale is not altered if the stimuli are spaced irregularly rather than at even intervals of intensity.

Procedure

The essentials of the experimental procedure were the same for all experiments reported here. One *S* at a time was seated in a soundproofed room with a headset comfortably adjusted on his head. All listening was monaural through a PDR-8 earphone mounted in a doughnut cushion. Every 6 sec. a 1000-cps tone was resented to him and he was required to make a single loudness rating for that tone on a score sheet which he held in his lap. He made his loudness judgments on a scale ranging from zero to 20, where "20" was to be used for the loudest tone heard in the series, "0" for a tone not heard at all, and "1" for the weakest audible tone. The other numbers were to be used for the intermediate loudnesses. The *S* was always specifically instructed to use the response categories as loudness categories. At no time did *S* know how many different tones were in the series, or what the intensities of the different tones were.

[1] This work was supported under Contract N5-ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This is Report No. 166-I-143, Project Designation No. NR-784-001 under that contract. The author is indebted to Miss Jeanne Jordan for assistance in the collection and tabulation of the data.

The tone lasted 1 sec. and was always preceded by the flash of a signal light lasting for 1 sec. and terminating 1 sec. before the onset of the tone. The entire pattern of light and tone was repeated once every 6 sec. The tone was filtered through a 1000-cps filter with a pass band of 100 cps, so that the onset and offset of the tone had few audible transients. The entire timing pattern of the tones and the light was controlled electronically, and *E* manually adjusted the intensity of the tone between each successive presentation.

Each different intensity was judged 100 times by each *S*. An experimental session lasted for 2 hr., with at least one and usually two rest breaks during the 2 hr. At the beginning of each experimental session, *S* was allowed to listen to the tones presented in random order until he felt that he had established the range of intensities and his judgmental criterion. Thus the data recorded do not reflect the initial period in which the *S* adjusts his judgments to the range of intensities presented.

All presented series were determined before running, and the series were constructed in such a way that at the end of each 400 presentations every stimulus intensity had preceded every other intensity an equal number of times. This precaution was taken to cancel out any possible successive contrast effects.

The *S*s were nine males and one female between the ages of 18 and 32. All *S*s had thresholds for the 1000-cps tone in the ear used of approximately 5 db SPL.

In treating the data each *S*'s data were handled separately to construct the equal discriminability scale. After each scale had been constructed, all scales were averaged together for the data presented here. It is necessary to treat each *S*'s data individually because all *S*s do not use the same judgmental criteria. Their discriminability scales may be very similar, but the actual judgmental categories used may be quite different. Thus pooling the data before determining how each *S* used the response categories can lead to serious distortion of the final scale.

Experimental Conditions

There were four different sets of experimental conditions used. The differences between them were in the range of intensities used and in the spacings of the intensities.

1. The range of intensities was 5 to 100 db re .0002 dyne/cm.^{2} (SPL) with stimuli spaced every 5 db. Thus there were 20 different intensities. Since each *S* judged each intensity 100 times, a total of 2000 judgments was made by each *S*. (For each condition, 100 judgments were made for each intensity.)

2. The stimuli were spaced every 5 db, but the intensity range was 55 to 100 db SPL.

3. The stimuli were spaced every 5 db, but the intensity range was 5 to 50 db SPL.

4. The range of intensities was 5 to 100 db SPL, but the stimuli were spaced roughly according to loudness rather than evenly on a decibel scale. Fractionation data had been obtained from the *S*s, and these fractionation data were used to construct a loudness scale. For these purposes, the main point is that the stimuli were spaced very unevenly in terms of log intensity. The acutal intensities used were: 5, 40.8, 55.6, 63.7, 69.3, 73.8, 77.5, 80.7, 83.4, 85.7, 87.8, 89.6, 91.3, 92.8, 94.2, 95.5, 96.7, 97.9, 99.0, and 100 db SPL. Thus the spacing of the intensities was radically different from that used in the other experiments.

The three series involving equal db steps of intensity were run in random order for each *S*. Only six *S*s were used for the last series, and this was run after *S* had completed the other series.

Results and Discussion

Although an intensity of 5 db SPL was used in three of the experimental series, no value for it could be determined because a few *S*s never heard that intensity. With no dispersion in the loudness judgments for that intensity, no scale value could be computed for it. However, the fact that that point was used experimentally should be remembered in some of the discussion.

Relation to cumulative DL's.—The open circles of Fig. 1 show the equal discriminability (E.D.) scale as a function of stimulus intensity when the entire range of intensities was used. Throughout the middle range of intensities (30 to 80 db), the function is essentially linear. The slope of the function is greater for lower and higher values, indicating that discriminability is relatively better at the higher and lower intensities.

If the relative DL for sound intensity were a constant, then a cumulative DL scale plotted against log intensity

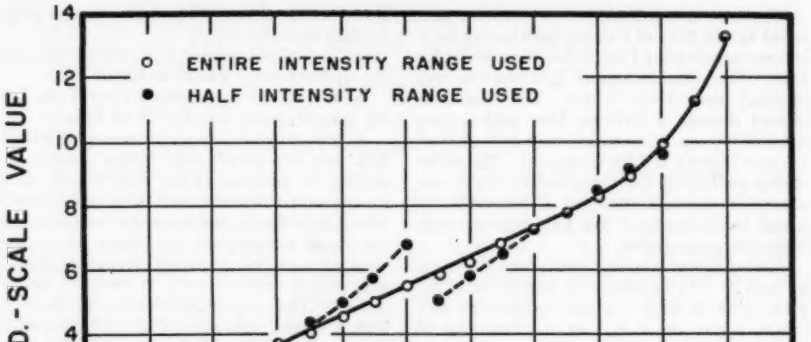

FIG. 1. A scale of equal discriminability as a function of tone intensity. For the open circles, the experiment was run with the entire range of intensities in one series. For the closed circles, the upper and lower intensity ranges were used in separate series. Since the intercept constant has no meaning on the E.D. scale, the curves have been anchored at a value of zero for an intensity of 10 db. The intercept constants of the three curves were adjusted independently to give best coincidence by a visual fit. Average data from 10 *Ss*.

would be a straight line, and we could conclude that in the middle range of intensities at least, the E.D. scale and a scale of cumulative DL's are linearly related. The data of Riesz (5), however, show that the relative DL for sound intensity rises as intensity is decreased; and a plot of cumulative DL's from the data of Riesz against the E.D. scale would not be a perfect straight line.

Harris *et al.* (4), however, have shown that the function relating the DL for white noise to sound intensity depends on the psychophysical method used. When they used a two-category method of constant stimuli (i.e., no indifference judgment permitted) they found that the relative DL for intensity was independent of intensity down to approximately 5 db Sensation Level. They argue that this method gives the best indication of the true sensitivity of the auditory mechanism, and that other methods allow a factor of "apprehension" to play too important a role. The forced-choice method is also most directly comparable to the technique used in the present experiments, since *Ss* were forced to make a scale judgment about every stimulus.

Although there are no data available which have used the forced-choice method with a 1000-cps tone, the results of Harris *et al.* allow us to assume that if such a method were used, the relative DL for intensity would be a constant regardless of intensity. In this case, the cumulative DL scale would be linearly related to log intensity and to the E.D. scale in the middle

range of intensities. The actual scale units for the two scales would be quite different, of course, but the linear relation makes it clear that the basic nature of the discriminative process is the same whether comparative or absolute judgments are made.

Anchoring effect of limiting intensities.—The linear relation between the E.D. scale and log intensity holds only for the middle range of intensities. At the higher and lower intensities, discrimination is relatively better. This effect we have presumed to be due to the anchoring produced by the intensity limits of the experimental series.

In order to test this assumption of the anchoring effect, the two additional experimental series were run with intensity limits corresponding to the middle of the entire intensity range used in the first series. If the better discrimination at the ends is due to an intensity limit effect, then the better discrimination should also occur in the middle intensities when these intensities are at the end of a series.

The data obtained from these two shorter-range series are also plotted in Fig. 1, and it is clear that the expected better discrimination does occur. Better discrimination occurred in the region of 50 db SPL when the two half intensities ranges were used than when the entire intensity range was used. The improved discrimination extends for about 15 to 20 db, which is about the same as the extent of the better discrimination at the two ends when the entire range was used.

Except in the region of 50 db, the curves show remarkable agreement. Thus the fact that a shorter range is used does not in itself make the function basically different. The difference in the curves occurs only at the ends of the functions.

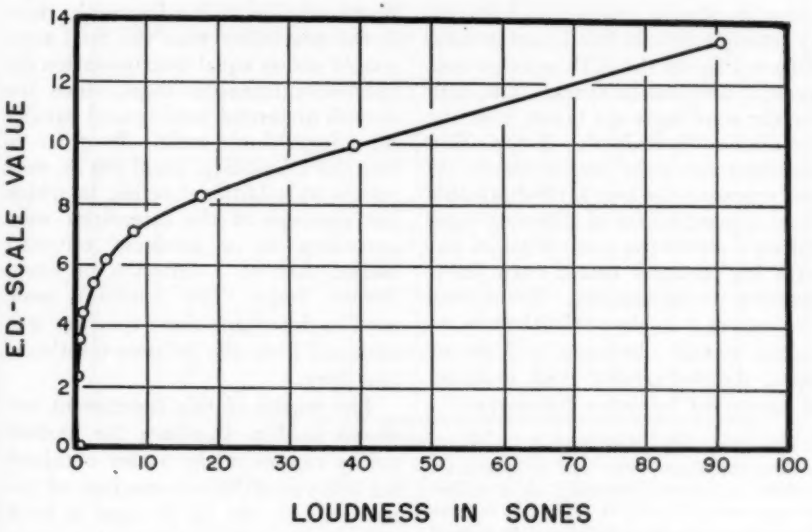

FIG. 2. The scale of equal discriminability as a function of loudness in sones. Loudness data are from American Standards Association (1).

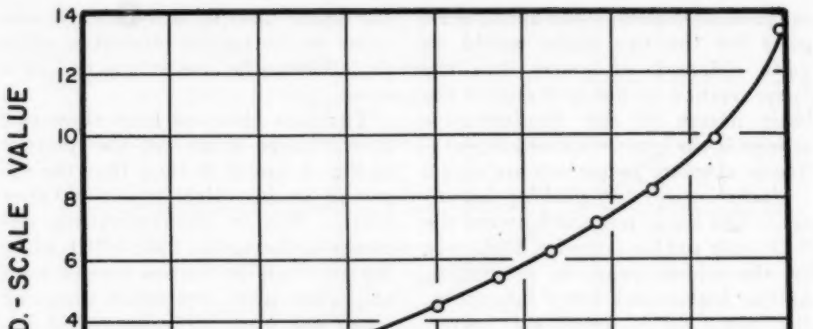

FIG. 3. The scale of equal discriminability as a function of log loudness. Loudness data same as Fig. 2.

Relation to loudness.—In contrast, the scale of equal discriminability bears no simple relation to loudness as determined from fractionation data. Figure 2 shows the E.D. scale plotted against loudness in sones. The data for the sone scale are taken from the American Standards Association standard for noise measurement (1) and represent the best scale obtainable from a compilation of different data. Figure 3 shows the same relation, but with log loudness rather than linear loudness on the abscissa. From these two curves it is clear that there is no simple relation between a scale of equal discriminability and loudness as measured by other techniques.[2]

[2] The fractionation data obtained from the *S*s in these experiments were used to construct a loudness scale also. Although this scale does not agree too well with the ASA scale, the conclusion about the relation of the scale of equal discriminability and the loudness scale is not changed if these data are used.

Effect of different spacings of intensities.—Since the E.D. scale showed a linear relation to log intensity, there is the possibility that the final scale simply shows equal discrimination for successive intensity steps, since the stimuli presented were spaced equally on a logarithmic scale. In order to test this possibility, six of the *S*s were run in an additional series, in which the spacings of the intensities were according to a loudness criterion rather than to a criterion of equal decibel steps. The loudness scale used to determine these spacings was obtained from the *S*s' own fractionation data.

The results of this experiment are shown in Fig. 4, where the plotted points represent the scales obtained for the two different spacings of intensities, for the six *S*s used in both experiments. It is clear that both experiments resulted in the same func-

tion, and that the shape of the E.D. scale is not a function of the spacing of the presented intensities.

A second point of interest from this comparison is that there is no secondary anchoring at approximately 40 db SPL. When stimuli were spaced by loudness, the lowest intensity used was 5 db, and the next lowest 40.8 db. With this wide separation, it seemed possible that the limit effect would occur near 40 db, rather than near 5 db. That no such anchoring occurred is obvious from the data. An interesting implication of this result is that it is possible to remove the anchoring effect from the discriminability scale by placing one stimulus well beyond the range of intensities for which the scale is desired. An anchor far enough removed from the presented points has no real anchoring effect itself, but it will prevent the limit-anchor from occurring at another point. In other words, the anchoring effect does not spread over a constant number of stimuli, but over a constant intensity range.

Information transmission.—In the previous paper by Garner and Hake (3) it was pointed out that there are many situations in which the stimulus is simply a symbol for an event or a class of events, and that if different stimuli are used to represent different possible events, it is meaningful to ask how much information a particular set of stimuli can transmit about the original events. It was further pointed out that the maximum amount of information would be transmitted if the stimuli are spaced according to a criterion of equal discriminability.

Although neither of the experimental series which used the entire

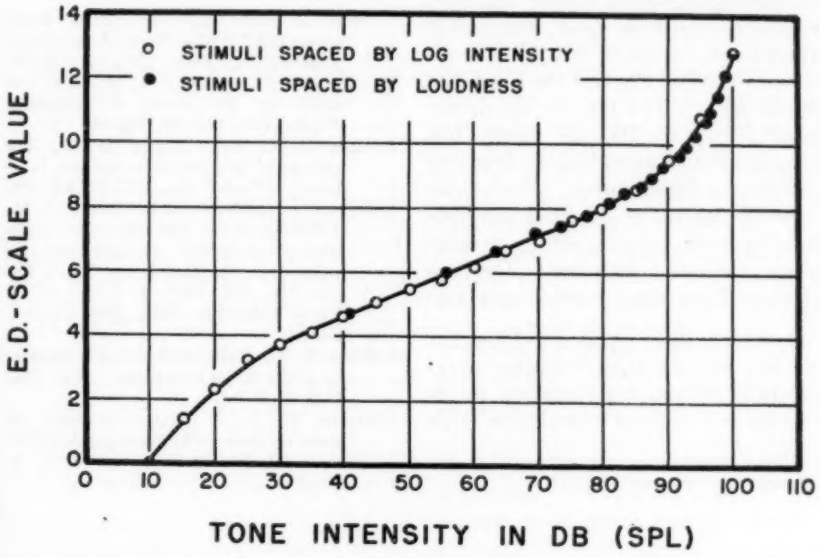

FIG. 4. The scale of equal discriminability as a function of tone intensity when presented stimuli were spaced every 5 db apart, and when they were spaced by a loudness criterion. Average data from six *Ss*.

intensity range used stimuli spaced exactly according to a criterion of equal discriminability, the series which used equal db spacings fulfills this criterion much more closely than the series which used equal loudness spacings. Thus an approximate test of this statement can be made with these data. For the six *S*s who participated in both of these experiments, the average amount of information transmitted in bits per stimulus presentation[3] was 1.68 for equal db spacings, and 1.48 for equal loudness spacings. Thus it is clear that greater information transmission will occur when the stimuli are spaced according to a criterion of equal discriminability.

Summary

Four experiments were run involving judgments of the loudness of tones presented one at a time. Three of the experiments spaced the presented stimuli 5 db apart, and one of these used a range of 5 to 100 db, another 55 to 100 db, and the third 5 to 50 db SPL. The fourth experiment used the 5 to 100 db range, but spaced stimuli according to loudness rather than log intensity. From these experiments scales of equal discriminability were constructed by a technique previously described. The conclusions from these experiments are:

1. The scale of equal discriminability constructed from loudness judgments bears the same relation to intensity as a scale of cumulative DL's for intensity. Thus the discriminative process is the same for absolute and comparative judgments.

2. Any experimental series of tones has definite limiting intensities at the high and low ends. Relatively better discrimination is shown near these limiting intensities, and it is suggested that a subjective anchor occurs at these points.

3. Spacing the presented stimuli non-uniformly on the intensity continuum does not change the scale of equal discriminability as long as the limiting intensities are the same.

4. The amount of information which stimuli can transmit about some other series of events will be greatest when the stimuli are spaced according to a criterion of equal discriminability.

(Manuscript received for priority publication December 17, 1951)

[3] For a description of this term, and its calculation, see Garner and Hake (3).

References

1. American Standards Association. Publication Z24.2-1942 on Noise Measurement. *J. acoust. Soc. Amer.*, 1942, **14**, 102–110.
2. Attneave, F. A method of graded dichotomies for the scaling of judgments. *Psychol. Rev.*, 1949, **56**, 334–340.
3. Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.
4. Harris, J. D., Myers, C. K., Stover, A. D., & Stuntz, S. E. The effect of sensation level on intensity discrimination for white noise. Medical Research Lab. Report No. 145, 1949, **8**, 1–26. (U. S. Naval Submarine Base, New London, Conn.)
5. Riesz, R. R. Differential intensity sensitivity of the ear for pure tones. *Phys. Rev.*, 1928, **31**, 867–875.
6. Saffir, M. A. A comparative study of scales constructed by three psychophysical methods. *Psychometrika*, 1937, **2**, 179–198.

ACCURACY OF TACTUAL DISCRIMINATION OF LETTERS, NUMERALS, AND GEOMETRIC FORMS [1]

T. R. AUSTIN AND R. B. SLEIGHT
The Johns Hopkins University

In problems of human efficiency there is a definite need for reducing visual and auditory burdens in an effort to relieve some of the strain placed on the human operator. This experiment was exploratory; it was an attempt to determine whether or not tactual cues might ultimately be employed in relieving some of this work load. Previous experimentation has found that tactual perception of form may be experienced most effectively by the tip of the tongue; however, the most likely applied situation involves finger-tip sensitivity which ranks second to the tongue in acuteness of tactual perception (**5**). For that reason this experiment was concerned only with the finger tip and specifically with the index finger tip.

Past experiments pertinent to this study may be broadly classified into three categories. First, there have been experiments determining the relative sensitivity of various parts of the body. In general, it may be said that the arm is more sensitive than the thigh (**2**), the palm more sensitive than the forearm (**3, 6**), the thumb and finger tip more sensitive than the palm (**3, 5, 6**), and the tongue the most sensitive of all (**3, 5**). Major (**3**) has studied tactual sensitivity of the lips but admits that finger-tip sensitivity is more effective.

Secondly, there have been experiments dealing with the adequacy of various stimulus materials. Calabresi (**1**) reports findings which indicate that elementary forms such as square, circle, and triangle are recognized more rapidly than more complicated forms. The latter necessitate reconstruction from knowledge of single parts. Major (**3**) tested two types of circles (solid and outline) and two types of triangles similarly constructed. He ranked the outline circles as easiest to discriminate and the solid circles as most difficult. The outline triangles and solid triangles ranked intermediate, in that order. Zigler and Barrett (**6**) employed five types of stimuli: squares, equilateral and right triangles, diamonds, and hexagons. Five sizes of each figure and three constructions (solid, outline, and point) were tested. The outline figures yielded the most accurate scores when tested on the thumb and the triangle ranked best of the figures tested. In another study, Zigler and Northrup (**7**) presented patterns of confusion found when several figures were tested. An analysis of such patterns would be critical before practical application of such figures would be possible.

The third category of the existing experimental literature involves study of discrimination through emphasis on the basic cutaneous experiences, viz., pressure, pain, warmth, and cold. Stone (**4**), for example, has had some success in studying form perception through the use of hot and cold stimuli.

From this very brief review of existing studies it seems apparent that

[1] This research was supported in part under the terms of a contract between Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This is report No. 166-I-146, Project Designation No. NR-784-001 under that contract.

information is lacking on the maximum number of forms which may be discriminated tactually; therefore, it was deemed worth while to study the problem of tactual sensitivity from the quantitative viewpoint.

The purpose of this experiment was, then, to determine to what extent various letters, numerals, and geometric forms could be discriminated tactually.

Materials

Two sets of stimulus figures were employed. One set contained all the letters in the alphabet, the numerals 2 through 9, and the following geometric forms or figures: diamond, square, right triangle, equilateral triangle, rectangle, parallelogram, crescent, semicircle, and cross. The figures comprising this set were cut from $\frac{1}{4}$-in. Masonite sheeting and each figure was the maximum size which could be inscribed in a $\frac{1}{2}$-in. circle. This seemed to be the maximum sensitive area utilizable on a plane surface by the average finger tip, as suggested by a study of fingerprints. Some slight distortions from common forms were made in the letters and numerals when such distortions enabled the over-all area to be increased (*cf.* Fig. 3). The letters and numerals had a constant stroke width of 2 mm.; the geometric forms were solid (had none of the inside area cut out).

The figures included in the second set were the same but were made up of wire pins set into

Fig. 1. Tactual discrimination apparatus. The disk containing the point figures is under the cover. The *S*'s hand is in position and the index finger is situated on a figure exposed through the slit in the cover. The other set of stimulus figures (solid) is shown in the background.

Masonite bases. While the external dimensions were the same as the first set, the stroke—instead of being a continuous line—appeared as a string of points outlining the figures. The points were 2 mm. apart; the stroke width in this set was 1 mm. The letters in both stimulus sets were upper case. In neither set were the numerals 0 (zero) or 1 tested because of the probable confusion with the letters O and I.

There was a total of 43 figures in each of the two sets. Each set of figures was mounted flatly on a 12-in. disk of plywood. The base of each figure paralleled the outside edge of the disk. The disks could be mounted under a cover and rotated by *E* in such a way that each figure was exposed separately through a slit in the cover. The cover could be removed and either of the disks used as desired. Figure 1 shows a photograph of the tactual discrimination apparatus.

Procedure

At no time during the experiment did *S* look at the figures he was to discriminate tactually. The apparatus was placed on a table in front of him and he was told that there would be three types of figures used: letters, numerals, and geometric forms. He was instructed to place the index finger on the figure at the given signal and to respond as quickly as possible by telling *E* exactly what the figure was. The *S* did not know before each trial whether the stimulus figure would be a letter, numeral, or geometric form; nor did he know that each figure would appear once and only once in each series. The time was taken and recorded for each response.

Both index fingers were used, changing hands every five figures on the point set and every ten figures on the solid set. This precaution was taken to reduce the effects of adaptation and fatigue. For both left- and right-hand positions, *S*s were told that the orientation of the figures was in line with the arm, not the body.

If any figure could not be discriminated, *S* could "give up" after a reasonable attempt, or was instructed to go on to the next figure if no discrimination was made within 1 min.

The procedure can conveniently be broken down into four parts, each part corresponding to a condition tested.

Condition 1: solid figures, no movement.—Employing the solid figure stimuli, *E* placed *S*'s finger tip on the stimulus figure. After the finger was so placed, no movement was allowed. The pressure could be varied slightly by a small amount of up and down or rolling movement, but at no time was the finger to be lifted from or moved on the figure. After a judgment had been obtained for this figure, *S* was instructed to lift his finger and keep his arm in place. The disk was revolved to the next position and the procedure repeated. Timing started the instant *E* released the finger and terminated at the first response from *S*. At this point, both the time and response were recorded on prepared data sheets.

Condition 2: solid figures, movement.—The solid figures were again used in this condition. Instead of *E* placing the finger on the figure, *S* probed until his finger was situated squarely on the figure. He was free to move or lift his finger, and could feel in any manner he chose to facilitate judgment. Timing began when the finger was square on the figure and ended when a response was made. As before, both the time lapse and response were recorded.

Condition 3: point figures, no movement.—The procedure for this condition paralleled that of the first condition except that the point figures were substituted for the solid figures. No movement was allowed, but the slight rolling or pressing was permitted. The finger was placed on the figure by *E* and responses and time periods were recorded.

Condition 4: point figures, movement.—The procedure of the second condition was repeated here with the point figures taking the place of the solid figures. The *S* placed his own finger on the figures and felt as he chose until a discrimination was made. Time and responses were again recorded.

Throughout the four conditions, the order of presentation of the figures in each condition was randomized so that letters, numerals, and geometric forms were presented in random sequence. A different random sequence was used for each *S* on each condition. Ten adult *S*s (five male and five female) were tested on each of the four conditions, hence a total of 40 random orders was employed.

Each *S* was tested on each condition only once. A retest would have made possible a measure of reliability but at the same time it would have introduced certain effects of familiarity which were not desired. The authors wished the results of this exploratory experimens to be as free from familiarity and learning at possible.

Results

Since a high degree of accuracy was the goal of this experiment, 90% accuracy was chosen as the acceptable criterion, i.e., not more than 10% error on a figure for the reports of all *S*s in any particular condition tested. The figures in each condition which

CONDITION	PERCENT ACCURACY								
	100	90	80	70	60	50	40	30	20
SOLID FIGURES NO MOVEMENT	I L T ▬	C - O F - 7 J - 2 O - 6 U - J	G - K, 6 X - E, Z Y - T, 8 ◆ - 8, ✱ 6 - ✎, ✱ ♭ - ▬, Å ❩ - C, 2	D - O, 9, ▣ E - L, ✱(2) P - F, 7, ✱ V - U, Y(2) Z - 8, Å, ✱ + - X(3)	A - R, Å(2), ✱ B - A, E, R, ✱ H - M, N, X(2) O - 8, O, 8, 9 3 - 2(2), 9, ✱ 4 - K, +, X, ✱ ▣ - ◆, ▬(3) ▰ - ▣(2), ▰, ✱	M - N, W(2), Å, ✱ N - H(2), M, W, 8 R - D, M, N, ✱(2) 9 - O, S, 5, 8, ✱ Å - A(2), 4, ▬, Å	K - E, F, X, Z(2), ✱ S - 8, E, F, 5, ✱(2) W - U(2), V(2), X, ✱ Z - 2, W, Z, 9, ✱(2) 5 - 8, C, E, S, X, ✱ 6 - C(3), G(2), 8 7 - T(5), ❣	8 - C, D, R, S(2), ▣, ✱	
SOLID FIGURES MOVEMENT	C G I J K L O Q U Y 7 9 ◆ ▬	D - O E - K P - R T - ◆ V - Y W - J X - Y 2 - Z ♭ - Å ❩ -) 6 - ◆	A - 4(2) F - E, ✱ R - A, K 5 - E, 5 Z - 2, 3 3 - 9,) 6 - 8, Q 8 - 8, ✱ ▣ - ▬, ▰ Å - ♭(2)	B - R(2), 8 M - N, K(2) 5 - S(2), Z + - X, ★(2)	N - H(2), W, X ▰ - ◆, ▣, Å(2)		H - B, K, M(2), W(2) 4 - K, Å, ★, ✱(3)		
POINT FIGURES NO MOVEMENT	I J L T U ▣	C - O F - Z O - C X - ✱ Y - K ♭ - ✱	K - ✱(2) R - K, X V - Y(2) ◆ - O, Q	A - X, ✱(2) D - O(2),) E - K(2), Z Z - ❩, ✱(2) + - X, ✱(2)	0 - 8, O, R, 6 H - W, X(3) P - F(2), 9, ✱ W - K, U(2), ✱ 2 - Z(3), 8 3 - 5, 9(2),) 9 - 8, 3, ▣, 6 Å - ♭, +, Å, ✱ ▬ - O, P, ▣, ✱ ▰ - ◆, ▣, Å, ▬	5 - E, 3, ▬, ✱(2) 6 - O, ◆, ▣, ❩, ✱	8 - D, M, P, R, 8, ✱ 5 - E(2), Z, 6, ◆, ▬ 8 - B(2), O, Å, ✱(2)	M - H(2), N, W, X(3) Q - O(3), 9, 6, ✱(2) 4 - A, M, Y, Å, ✱(3) 6 - E(2), G(2), R, 5, ▣ 7 - K, T(5), ✱ ❩ - D, Z, 3(2), 9, ✱(2)	N - H(3), K, X(3), ✱
POINT FIGURES MOVEMENT	C I L O U	A - 4 E - K G - O H - N K - X R - A T - X V - Y X - ★ 3 - 9 Å - Å ♭ - Q	F - K, Y J - U, 6 P - R, ♭ Y - V, X 8 - 6, ▣ 9 - Q, 8	D - O(2), ✱ Z - X, ★, ✱ ◆ - O(2), ▣ ▣ - O(2), ✱ ▬ - ▣(2), ▰ ▰ - 8, ▬, X 6 - O(3) + - K, ✱(2)	N - K(2), X(2) S - 5, 9, ▣, ✱ W - H, ✱(3) 2 - R, Z, ✱(2) 7 - T(3), ✱	8 - O, 8(4) 5 - E, S(2), 6, ▣	M - H(4), N, ✱ Q - R(2), 2, ▣, 6, ✱ 4 - A, +, Å, ✱(3) ❩ - C, 2, 3, 9, ✱(2)	6 - G(4), S, 8, ✱	

FIG. 2. Per cent accuracy for all figures tested on each of the four conditions. The character on the left in each column is the stimulus; those to the right of it are the incorrect reports. Asterisks denote errors made by *S* "giving up." Numbers in parentheses indicate number of times preceding error was made.

met this criterion will be emphasized in the remainder of this report. Results will be divided into four parts corresponding to the four conditions of the experiment.

Condition 1: solid figures, no movement.—Figure 2 shows the complete error pattern for each test condition. In each per cent accuracy column are found the figures reaching the respective accuracy levels and the errors made on those figures. Errors accounted for by *S*s giving up are shown by asterisks in the figure.

In this condition an accuracy of 100% was attained for the I, L, T, and rectangle; the C, F, J, O, and U met the 90% accuracy criterion. In all, there were nine figures which met the criterion. The responses for any specific figure may be seen by reading from left to right in each per cent accuracy column. For example, there were no errors on the I, one error on the C (given as O), two errors on the G (given as K and U), three errors on the D (given as Q, 9, and square), etc.

Condition 2: solid figures, movement.—In this condition 25 figures met the 90% or better accuracy criterion. Fourteen figures (C, G, I, J, K, L, O, Q, U, Y, 7, 9, diamond, and rectangle) were never reported incorrectly; and 11 figures (D, E, P, T, V, W, X, 2, right triangle, crescent, and semicircle) reached the 90% accuracy level. All responses are again recorded in Fig. 2 and the figure is interpreted in the same manner as explained above.

It will later be shown that this condition (solid figures, movement) yielded the highest degree of accuracy of all conditions tested.

Condition 3: point figures, no movement.—Twelve figures were discriminated with an accuracy of 90% or better in this condition; the I, J, L, T, U, and square at the 100% level, and the C, F, O, X, Y, and right triangle at the 90% level. Figure 2 shows that one *S* gave up (made no discrimination) on the X and the right triangle stimuli.

Condition 4: point figures, movement.—The results of the fourth condition are recorded in the last row of Fig. 2. Seventeen figures met the criterion; the C, I, L, O, and U at the 100% accuracy level, and the A, E, G, H, K, R, T, V, X, 3, right triangle and equilateral triangle at the 90% level.

Discussion

Over-all evaluation.—Figure 2, which has been discussed in detail above, may be considered the over-all response pattern. It is interesting to note that the I and L were reported correctly throughout the experiment. Figures common to all four conditions which met the criterion of 90% accuracy or above are: C, I, L, O, T, and U. These figures, therefore, are easy to discriminate tactually no matter how they are presented.

In the applied situation where for various reasons, e.g., uniformity, it is desirable to use a set of figures all constructed in the same fashion, one might wish to know which of the tested conditions should be utilized to convey information the most effectively. Obviously, when no movement is permitted, not as much information can be gained through the finger tip as when movement is allowed. With no movement of the fingers, 9 solid and 12 point figures met the criterion. When movement was allowed, however, 17 point and 25 solid figures were discriminated with 90% or better accuracy. In the solid figures—movement condition, almost 60% of the stimulus items met the criterion. These results seem quite remarkable considering the rigid ex-

FIG. 2. Per cent accuracy for all figures tested on each of the four conditions. The character on the left in each column is the stimulus; those to the right of it are the incorrect reports. Asterisks denote errors made by *S* "giving up." Numbers in parentheses indicate number of times preceding error was made.

perimental conditions which included no indoctrination of *S*s and a 1-min. maximum discrimination time. The sizes and shapes of the 25 figures in the solid figure—movement condition which met the 90% criterion are drawn to scale in Fig. 3.

Time factor.—It was thought at the beginning of the experiment that the items which were discriminated the most accurately might also be discriminated fastest. For this reason records were kept for the time of each response. A product-moment correlation coefficient was computed on the total data between the number of times each figure was discriminated correctly and the speed of correct responses. The obtained r was $-.14 \pm .15$, a negative relationship between the number of correct responses and length of discrimination time. This correlation coefficient was not significantly different from zero, showing for these total data that speed of response and correctness of discrimination are not related.

Sources of error.—There were two general sources of error in the judgments. First, there was the "give-up" category. If *S* could not respond to a stimulus (although he was encouraged

FIG. 3. Exact sizes and shapes of the 25 stimulus items which met the criteria of 90% or better accuracy in Cond. 2 (solid figures, movement)

to guess in such cases), or if he did not respond within 1 min., the would-be response fell into this category. Not many errors were accounted for in this manner; wherever they occurred they seemed to be caused by a lack of sensory cues on the more finely detailed figures.

The second type of error occurred when one figure was confused with another. This is best illustrated by the confusions which occurred among H, N, M, and K. Examples of confused pairs are: T and 7, Z and 2, G and 6, B and 8, A and equilateral triangle, C and 0, S and 5, etc.

The second source of error suggests a way of making the series more discriminable. When two figures are often confused with each other, the elimination of one of the pair should raise the accuracy score of the second. The T and 7, and B and 8 will serve to illustrate this point. The 7 was often called a T but the T was seldom called a 7. In this case, elimination of the 7 would but slightly aid discrimination of the T, whereas elimination of the T might greatly enhance discrimination of the 7. In the final analysis, elimination would depend upon which one of the pair was the most critical to a specific problem.

The B and 8 were confused with each other with approximately equal frequency. Elimination of either should enhance discrimination of the other.

Application.—It might be well at this point to consider some of the possible applications of tactual cues. The zero correlation noted earlier between accuracy and speed of response for all the data taken together does not mean that the time factor is unimportant when considering the best discriminated figures. Let us assume that ten figures are needed for tactual facilitation on a keyboard. Which ten figures should be employed? It may further be assumed that these necessary figures will be drawn from the solid figure set since this set yielded the highest accuracy score. Three possible criteria may be utilized in the selection of the ten figures: speed, accuracy, and a combination of both.

The 25 figures may be ranked by accuracy and then the first 10 selected. The first column in Table 1 shows ranking for accuracy. If this condition were used, the following figures would be selected for the keyboard: C, I, O, 7, L, U, J, diamond, 9, and K.

However, it may be more critical to set speed as the criterion as long as accuracy is limited to the minimum of 90%. The 25 figures may be ranked by speed and the first ten again selected. According to Table 1 the following would be selected: C, I, V, O, 7, L, U, D, E, and T. The reader may notice that six out of ten are common to both selections: C, I, O, 7, L, and U.

To set the most effective criterion, both speed and accuracy should most likely be considered. To select by accuracy alone involves using items with much longer reaction times than necessary, and selection by speed alone involves too many inaccuracies. For these reasons it seems safer to combine the two criteria. The third column in Table 1 illustrates a ranking by speed and accuracy combined. In making such a ranking it is necessary to give an arbitrary weighting to those items meeting the 90% accuracy criterion when comparing them with the 100% accurate items. This has been done by assigning an extra time unit (1 sec.) to the 90% accurate items and then ranking for time. This is a rough weighting procedure but it will suffice for the purpose of putting in a more favorable light the items which were responded to rapidly but had slight inaccuracy. A formal weighting procedure was not justified because of the limited nature of the accuracy data, skewness of the time data and use of a "give-up" category in the accuracy data.

Employing the combined speed and accuracy criterion, the following figures would be selected: C, I, V, O, 7, L, U, J, D, and E. This set of figures has seven items in common with criterion one and nine figures in common with criterion two. The fact that only six out of ten items (C, I, O, 7, L, and U) are common to all three criteria suggests that each specific applied situation would be improved by selection of a valid criterion before choosing the figures to be employed.

At present it is impossible to make a definite statement concerning the application of tactual cues, but the possibilities are numerous. Tactual cues, given to the finger tip by means of raised numerals, letters, and geometric forms, might increase use of control of instrument panels or keyboards. In general, facilitation might occur in any situation which involves the use of the finger tips in the performance of tasks where visual cues are not readily available or are impossible. This report covers only a small

TABLE 1

TWENTY-FIVE FIGURES MEETING THE 90% ACCURACY CRITERION FOR THE SOLID FIGURES-MOVEMENT CONDITION RANKED FOR ACCURACY, SPEED, AND A COMBINATION OF BOTH
Each recorded time represents a mean of ten subjects.

Ranking by Accuracy				Ranking by Speed		Ranking by Speed and Accuracy Combined	
Rank	Figure	Mean Time in Sec.	% Accuracy	Rank	Figure	Rank	Figure
1	C	4.6	100	1	C	1	C
2	I	5.3	100	2	I	2	I
3	O	6.5	100	3	V	3	V
4	7	6.9	100	4	O	4	O
5	L	7.0	100	5	7	5	7
6	U	7.6	100	6	L	6	L
7	J	8.1	100	7	U	7	U
8	Diamond	9.8	100	8	D*	8	J
9	9	9.9	100	9	E*	9	D*
10	K	10.4	100	10	T	10	E*
11	Rectangle	11.3	100	11	J	11	T
12	G	14.0	100	12	W	12	Diamond
13	Y	14.4	100	13	Diamond	13	9
14	Q	18.4	100	14	9	14	K*
15	V	5.4	90	15	K	15	W*
16	D*	7.7	90	16	P	16	Rectangle
17	E*	7.7	90	17	X	17	P
18	T	8.0	90	18	Rectangle*	18	X
19	W	9.4	90	19	Rt. triangle*	19	Rt. triangle
20	P	10.6	90	20	G	20	G
21	X	10.9	90	21	Semicircle	21	Semicircle
22	Rt. triangle	11.3	90	22	Y	22	Y
23	Semicircle	14.3	90	23	Crescent	23	Crescent
24	Crescent	14.7	90	24	2	24	2
25	2	15.7	90	25	Q	25	Q

* Both members of respective pair are equivalent and may be interchanged.

range of figures; many others might be studied and found as good or better than those tested.

Possible specific applications of tactual cues are: typewriter keyboards to aid learning of the touch system, push buttons on industrial machinery in cases in which the operator's eyes are occupied elsewhere, cockpit controls to relieve the visual and auditory strain on the pilot, and dashboard switches on the ordinary automobile. In other words, tactual cues may speed up learning or efficiency, or relieve the burden on other sense modalities and thus promote better overall human efficiency.

SUMMARY

The purpose of this experiment was to determine to what extent various letters, numerals, and geometric forms could be discriminated tactually with the index finger tip. Two sets of stimulus figures were tested, one set of solid figures and one set which contained figures made up of points. Both sets included all the letters in the alphabet, the numerals 2 through 9, and various geometric forms. The stimulus figures were presented one at a time in random sequence to ten *S*s.

The experiment involved four conditions of presentation of the stimulus figures: (1) solid figures, no finger-tip movement permitted; (2) solid figures, movement; (3) point figures, no movement; and (4) point figures, movement.

Of the four conditions tested, the solid figures—movement condition proved to be the most effective. Un-

der the restrictions of this condition, approximately 60% of the figures were discriminable at or above the 90% level of accuracy, i.e., not more than one error from a total of ten *Ss* on each of the figures. These figures were the following: C, D, E, G, I, J, K, L, O, P, Q, T, U, V, W, X, Y, 2, 7, 9, diamond, right triangle, rectangle, crescent, and semicircle.

No errors were made on the I or L throughout the experiment regardless of mode of presentation, and the C, I, L, O, T, and U met the 90% accuracy criterion in each of the four conditions.

Possible applications of tactual cues were suggested and three criteria for selecting the most effective figures were discussed.

(Manuscript received for priority publication December 11, 1951)

References

1. Calabresi, R. La percezione tattilo-cinetica delle grandezze e delle forme negli adulti e nei fanciulli. *Riv. di sci. appl. all'educ. fis.*, 1931, **2**, 1–37. (Abstract)
2. Franz, S. I. Diffusion effects following localized tactile training. *Publ. Univ. Calif. Los Angeles, Educ., Phil., Psychol.*, 1933, **1**, 129–135.
3. Major, D. R. Cutaneous perception of form. *Amer. J. Psychol.*, 1898, **10**, 143–147.
4. Stone, L. J. An experimental study of form perception in the thermal senses. *Psychol. Rec.*, 1937, **1**, 235–337.
5. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.
6. Zigler, M. J., & Barrett, R. A further contribution to the tactual perception of form. *J. exp. Psychol.*, 1927, **10**, 184–192.
7. Zigler, M. J., & Northrup, K. M. The tactual perception of form. *Amer. J. Psychol.*, 1926, **37**, 391–397.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	–	4	5	–	$1.25	$5.00
1920	3	–	2	–	4	5	–	$1.25	$3.75
1921	4	1	–	3	4	5	6	$1.25	$6.25
1922	5	–	–	3	4	5	–	$1.25	$3.75
1923	6	–	–	–	–	–	–	——	——
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	2	3	4	–	6	$1.25	$6.25
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	–	–	3	4	5	6	$1.25	$5.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	1	2	3	4	5	6	$1.25	$7.00
	27	1	2	3	4	5	–	$1.25	$6.25
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	1	2	3	4	5	6	$1.25	$7.00
1949	39	1	2	3	4	5	6	$1.25	$7.00
1950	40	1	2	3	4	5	6	$1.25	$7.00
1951	41	1	2	3	4	5	6	$1.25	$7.00
	42	1	2	3	4	5	6	$1.25	$7.00
1952	43	By subscription, $14.00 (2 vols.)						$1.25	——
	44								
		List price, Volumes 1 through 42							$269.00
		30% discount							80.70
		Net price, Volumes 1 through 42							$188.30

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year; from 1944 to 1950, in one volume per year. Beginning in 1951 there are two volumes a year.

Information about prices: when all numbers are available, the price per volume throughout the series is $7.00. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Address orders and subscriptions to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue N.W. Washington 5, D. C.